

# Cell

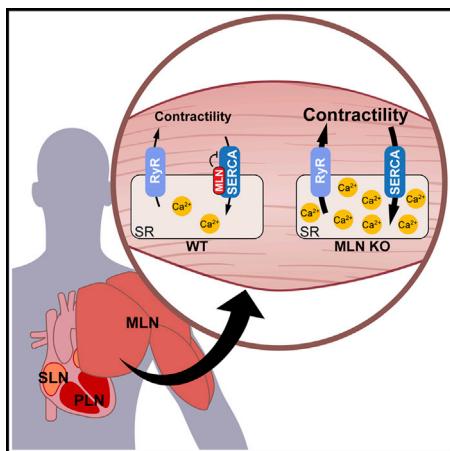
Volume 160  
Number 4

February 12, 2015

[www.cell.com](http://www.cell.com)



**Strain-Level Variation  
in the Gut Microbiome**



## Micropeptide's Brawny Powers

PAGE 595

Anderson et al. identify a skeletal muscle-specific micropeptide called myoregulin, encoded within a long noncoding RNA, that controls skeletal muscle function by directly inhibiting the pump activity of the  $\text{Ca}^{2+}$ -ATPase SERCA. Mice lacking myoregulin demonstrate improved intracellular calcium handling and exercise performance.

## Gut-Level Individuality

PAGE 583

The species composition of gut microbiomes can differ markedly between individuals. Greenblum et al. now uncover extensive variation within individual gut microbial species in their gene composition and copy number and link these strain-level differences to obesity and inflammatory bowel disease.

## Decoding a Long Noncoding Mystery

PAGE 607

EBER2 is a highly abundant noncoding RNA with unknown function that is expressed by the Epstein-Barr virus. Lee et al. find that EBER2 localizes to specific repeat sites on viral chromatin to facilitate binding of its interacting host transcription factor PAX5 to these sites. Recruitment of the EBER2-PAX5 complex depends upon RNA-RNA interactions between EBER2 and nascent transcripts that together control the viral lytic cycle and infection.

## An Extracellular Stop to Influenza

PAGE 631

Interferon-stimulated genes (ISGs) act in concert to provide a tight barrier against viruses. Using a screen to identify ISGs inhibiting late stages of the influenza A virus infection, Dittman et al. find that plasminogen activator inhibitor (PAI-1) blocks maturation of the viral surface glycoprotein, thus reducing virus spread in the airways. These findings show that the innate immune system, driven by type I interferon, uses modulation of the extracellular environment to inhibit viruses.

## Stress Test for RNA

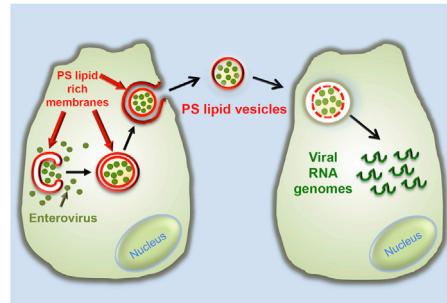
PAGE 644

The CCA-adding enzyme monitors the stability of tRNAs and tRNA-like small RNAs. Whereas CCA is added to stable RNAs, CCACCA is added to unstable ones to initiate their degradation. Kuhn et al. now characterize how these two scenarios are distinguished. Following CCA addition, stable RNAs are ejected, whereas unstable RNAs refold and are subjected to a second round of addition. Therefore, RNAs proofread themselves through differential responses to the interrogation of the enzyme.

## Switching Speed with a Squeeze

PAGE 659 and PAGE 673

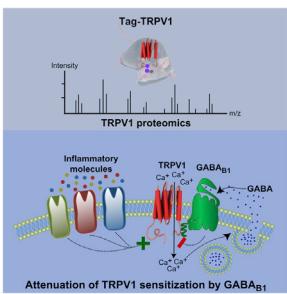
Fast amoeboid cell migration in 3D environments is central to developmental and disease-related processes, such as cancer metastasis. Liu et al. and Ruprecht et al. now find that changes in the environment, namely absence focal adhesion and strong confinement, trigger the switch in motility behavior. Liu et al. analyze a large set of different slow mesenchymal cells and show that the requirements for the motility switch are universal. Ruprecht et al. demonstrate that *in vivo*, during zebrafish development, fluctuations in cortical contraction forces act in concert with confinement and low adhesiveness, allowing embryonic progenitor cells to acquire a fast and persistent migratory behavior.



## Viruses Gang Up

PAGE 619

A central paradigm in virology is that each virus largely behaves as an independent infectious unit. Chen et al. now demonstrate that clusters of enteroviruses are packaged within phosphatidylserine (PS)-enriched vesicles that are nonlytically released from cells. Viruses within these vesicles have enhanced infection efficiency, and PS lipids serve as cofactors for enterovirus receptors. Clustering viruses within vesicles enables multiple viral genomes to be collectively transferred into cells, facilitating genetic cooperativity among viral quasispecies and promoting viral replication.



## Curbing Pain

PAGE 759

Inflammatory pathways stimulate the capsaicin receptor TRPV1, leading to pain hypersensitivity. Hanack et al. identify a feedback mechanism that keeps TRPV1-mediated pain in check and could lead to new avenues for analgesic therapy. Release of the neurotransmitter GABA in response to TRPV1 activity triggers a noncanonical signaling pathway that inhibits only hyperactive TRPV1, leaving homeostatic pain responses intact.

## Catastrophic Cure

PAGE 686

Chromothripsy is a catastrophic cellular event, first described in cancer, in which chromosomes are shattered and pieced back together imperfectly. McDermott et al. report a remarkable case in which a woman with WHIM syndrome, an inherited immunodeficiency disease, was fortuitously cured by a chromothriptic event in a single hematopoietic stem cell that deleted the abnormal copy of the disease gene. This cell then took over the bone marrow and restored normal immune function.

## Complement-ing Tumor Suppression

PAGE 700

Macrophages have a crucial role in mediating inflammation that contributes to development of cancer. Bonavita et al. report that deficiency of PTX3, a soluble mediator that regulates Complement activity, is associated with cancer in humans and causes tumor-promoting macrophage recruitment, angiogenesis, and *Trp53* mutations in mice. Thus, PTX3 acts as a physiological extrinsic oncosuppressor inhibiting tumor generation through modulation of the activity of macrophages.

## A Malignant Energy Switch

PAGE 715

AMPK is the master regulator of cellular energy homeostasis, with tumor suppressive activity. Pineda et al. describe a widespread mechanism by which AMPK is inactivated in cancer. Cancer-specific MAGE-A3/6-TRIM28 E3 ubiquitin ligase is an oncogenic driver that ubiquitinates and degrades AMPK $\alpha$ 1, resulting in downregulation of autophagy and increased mTOR signaling. These findings identify the mechanism of action of the MAGE-A3/6 cancer-testis antigens and illustrate a regulatory axis for altering cellular metabolism in cancer.

## Chaotic Therapy for Cancer

PAGE 729

The ERK kinase has been long thought to be the only substrate of MEK. Tang et al. now find that HSF1, the master regulator of proteotoxic stress responses, is a new MEK substrate. MEK blockade inactivates HSF1 and provokes proteomic chaos. Tumor cells are particularly susceptible to proteostasis disruption. In fact, amyloidogenesis induced by MEK inhibition suppresses tumor growth, suggesting that disruptions of the fragile tumor proteostasis may be feasible as therapeutic strategy.

## Inflaming Metabolic Dysfunction

PAGE 745

A high-fat diet is associated with several metabolic abnormalities such as hyperglycemia and increased rates of hepatic glucose production. Perry et al. demonstrate that these abnormalities are primarily caused by increased macrophage-derived IL-6 levels in white adipose tissue that results in impaired lipolysis and a concomitant increase in acetyl coA levels that fuel hepatic glucose production. Inflammation, thus, is likely the underlying trigger of metabolic dysfunction associated with a high-fat diet.

## Constant Change Is the Mother of All Frustration

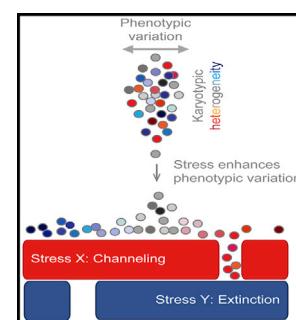
PAGE 785

Pathogens such as HIV and influenza are highly mutable. Their neutralization requires the generation of antibodies that cross-react with different viral strains, which it is hard to achieve. Using an in silico approach, Wang et al. find that cross-reactive antibodies occur with low probability because conflicting selection forces, imposed by the presence of different variants of the same antigen, frustrate affinity maturation. Importantly, frustration can be overcome by sequential immunization approaches.

## Conformity Quells Resistance

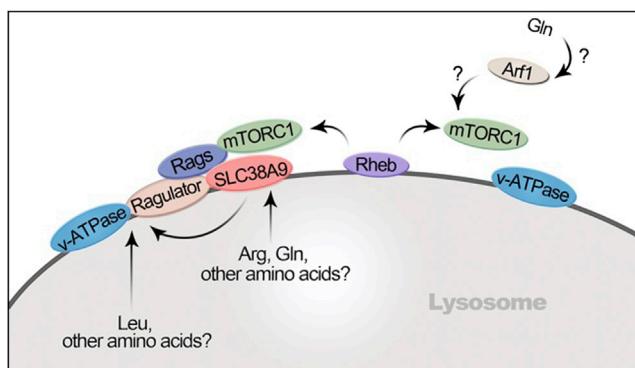
PAGE 771

Many diseases that readily evolve drug resistance are caused by cell populations that acquire diverse karyotypes or chromosome copy number. Chen et al. find that growth-suppressing stresses only serve to increase the heterogeneity of aneuploid populations, causing resistance to emerge. Using a fungal pathogen as a model system, the authors propose and test a strategy termed an “evolutionary trap” to eradicate such populations. One stress is applied to homogenize the population via adaptation; a second specifically targets and eliminates the new newly dominant karyotype.



# Amino Acid Intricacy

As the central bridge connecting nutrient sensing and cellular activities, mTORC1 has for years drawn much attention in the cell metabolism field. Its frequent dysregulation in cancer and the diverse therapeutic potential of the mTOR inhibitor rapamycin have further spread its acclaim among cancer biologists, clinicians, and pharmaceutical companies. Although efforts have yielded an in-depth understanding of the molecular components that regulate mTORC1 localization and activity at various steps along the signaling cascade, a key issue that has not been adequately addressed is the amino acid specificity in mTORC1 regulation. Does every flavor of amino acid taste the same to mTORC1? If not, how are they distinguished, and perhaps more importantly, why does it matter?



Amino Acid Sensing for mTORC1 Regulation. Courtesy of K.L. Guan.

A recent study from the group of Kun-Liang Guan makes a breakthrough in pursing this line of inquiry (Jewell et al., 2015). RagA and RagB are well-characterized essential regulators for the lysosomal recruitment and activation of mTORC1. Surprisingly, instead of paralyzing the entire mTORC1 pathway, they find that RagA/B deficiency in mouse fibroblasts selectively impairs the ability of mTORC1 in sensing leucine and arginine, while its sensitivity to glutamine remains intact. Together with an earlier work from Michael Hall and colleagues in yeast (Stracka et al., 2014), the findings highlight the mechanistic differences in TORC1 activation upon stimulation of different amino acids. Further investigation reveals the presence of an ADP ribosylation factor 1 (Arf1)-dependent, Rag-independent pathway for mTORC1 recruitment to the lysosome, suggesting a context-specific role, rather than universal requirement, of Rag in mTORC1 localization. Although this pathway appears to be dedicated to glutamine-mediated mTORC1 regulation and leads to distinct recruitment kinetics compared to its Rag-dependent counterpart, the relative contribution of these two pathways to mTORC1 activation in normal physiological context remains to be determined. As Jewell et al. (2015) mention in their paper, glutamine addiction is a feature shared by a variety of cancer cells. Therefore, it would be intriguing to explore whether this glutamine-specific sensing mechanism would be particularly important in the cancer

setting and to test its potential as a target for anti-cancer therapeutic development.

The work in yeast and mouse fibroblasts reinforces the notion that TORC1 does not see every amino acid in the same way and suggests that the mechanisms behind this may be evolutionary conserved. However, this leaves unresolved the identity of the direct amino acid sensor. Two studies from Giulio Superti-Furga and David Sabatini represent the most recent progress toward this goal (Rebsamen et al., 2015; Wang et al., 2015). The two groups independently characterize a member of the solute carrier family 38, SLC38A9, as a lysosomal membrane-resident amino acid transporter that is required for mTORC1 activation by certain amino acids. The characterization of the amino acid binding and transportation activity of SLC38A9 is exciting and certainly in line with the putative role of SLC38A9 being an amino acid sensor. However, fully supporting this notion requires direct evidence linking amino acid binding and its impact on signal transduction. Moreover, it appears that SLC38A9 has a rather non-specific substrate profile. It is therefore unclear whether, and if so to what extent, SLC38A9 might be involved in differential mTORC1 regulation by various amino acid inputs. In addition to lysosome membrane-based sensors, the work from Jewell et al. (2015) and others suggests that differential sensing occurs even before mTORC1 recruitment to the lysosome. The molecular basis for this early-stage sensing is yet another open question. Indeed, as much as we have already known about mTORC1 regulators, we are just getting started in appreciating the mechanistic intricacy of the amino-acid-sensing component of mTORC1 signaling.

## REFERENCES

Jewell, J.L., Kim, Y.C., Russell, R.C., Yu, F.X., Park, H.W., Plouffe, S.W., Tagliabracci, V.S., and Guan, K.L. (2015). *Science* 347, 194–198.

Rebsamen, M., Pochini, L., Stasyk, T., de Araújo, M.E., Galluccio, M., Kandasamy, R.K., Snijder, B., Fauster, A., Rudashevskaya, E.L., Bruckner, M., et al. (2015). *Nature*. Published online January 7, 2015. <http://dx.doi.org/10.1038/nature14107>.

Stracka, D., Jozefczuk, S., Rudroff, F., Sauer, U., and Hall, M.N. (2014). *J. Biol. Chem.* 289, 25010–25020.

Wang, S., Tsun, Z.Y., Wolfson, R.L., Shen, K., Wyant, G.A., Plovanich, M.E., Yuan, E.D., Jones, T.D., Chantranupong, L., Comb, W., et al. (2015). *Science* 347, 188–194.

Jiaying Tan

## Mary Frances Lyon (1925–2014)

Mary Lyon was one of the most notable geneticists of the 20th century. She is renowned for her discovery of X inactivation, an early example of epigenetic gene regulation, but she also made fundamental contributions to the entire field of genetics.

Mary was born on May 15, 1925 in Norwich, in the rural east of the United Kingdom. She was the first child of Clifford James Lyon, a civil servant, and Louise Frances Lyon (nee Kirby), a former school teacher. When she was 10, Mary's family, which by then included a younger brother and sister, moved to Birmingham. There, she attended King Edward VI High School for Girls and began her scientific career, like so many of us, through an inspirational teacher who made lessons interesting: Mary Udall, whom Mary described as having "a clear analytical mind." Mary was fascinated by physics and chemistry, but the prize she received from an essay competition, four books on nature study, ultimately led to her becoming a world-class biologist.

The Second World War helped change the status of women, and Mary decided to take the unusual step of reading natural sciences at the University of Cambridge (Girton College), where women took the same coursework as men but at that time were awarded only titular degrees. She focused on zoology and became interested in experimental embryology, graduating in 1946 with her titular degree. Rather unusually, she went on to Ph.D studies with R.A. Fisher, the Balfour Chair of Genetics at Cambridge, who founded much of the field of statistics and developed analysis methods for early gene linkage studies.

R.A. Fisher was in his fifties—brilliant, eccentric, and difficult. He threw out many who joined his lab, but he led Mary directly to her lifelong study of mouse genetics. At the time, there was no systematic method of mapping mouse genes. With 20 mouse chromosomes, Fisher decided that, if he crossed all possible combinations of his 21 visible mouse mutants, he would likely detect a new linkage. Students were each given a line of mice carrying five mutations, and Mary took on line 18, which included the

"pallid" mutation. Already an insightful experimentalist, Mary was unconvinced that her mouse-crossing experiments would give her enough data for a Ph.D, but she noticed that pallid mice tended to tip their heads to one side. She found that they were missing otoliths in the inner ear and went on to investigate the effects of penetrance and to work out how otolith absence correlated with postural reflexes.

Meanwhile, Mary was reading books by Conrad Waddington, known as the father of epigenetics, and became aware of the novel idea that embryonic development depends on genes—at a time when the exact nature of a gene was unknown. R.A. Fisher was not interested in developmental genetics, and Mary needed facilities for histology; so as Waddington had returned from his wartime post (scientific advisor to the Royal Air Force) to become Professor of Animal Genetics at the University of Edinburgh, Mary moved to his department and to a new supervisor, Douglas Falconer.

Waddington and Falconer were important influences on Mary; she chose to stay in mouse genetics due to the relevance to human studies. Waddington applied to the Medical Research Council (MRC)—which unlike other funders awarded equal pay for women—and gained postdoctoral funding for Mary to

stay in Edinburgh with Toby Carter, working on the inherited genetic risks from exposure to ionizing radiation, an area of great concern after WWII. Several novel mutants came out of this research, but facilities were not available for the amount of mouse breeding required, so in 1955, the entire mouse group moved to the MRC Radiobiological Research Unit at Harwell, led by John Loutit. There, Mary worked for an astonishingly productive period spanning more than 50 years. In the late 1950s, much of her research centered on the novel chromosome translocations from the Edinburgh mutagenesis work, working with Carter, Tony Searle, and one of the early great cytogeneticists, Charles Ford.

Mary knew about X-linked mouse mutants since her Edinburgh days, and at Harwell she focused on mottled mutants, in which females have patches of two different coat colors. She noted that males either died in embryogenesis or had a single coat color. However, one mouse grabbed her attention (ironically, a spontaneous mutant, not a radiation mutant): a mottled male. Mary bred this male and realized that, if a mutation had occurred when he was an embryo of just a few cells, he would be a mosaic of mutated and normal X chromosomes. Her analysis led to the discovery that this could also apply to his mottled daughters, who had two types of cells: one with an active gene and one without.

Mary's knowledge that female mice (unlike female humans) need only one X chromosome for normal development and have sex chromatin in their nuclei led her straight to the work of Susumu Ohno, whom she regarded as an exceptionally creative and gifted scientist, and his discovery that the sex chromatin was a condensed X chromosome. Mary had already considered the idea that only one X chromosome is active in females, but her analysis of the mottled mouse, in the light of Ohno's discovery, led her to publish her X inactivation hypothesis in *Nature* in 1961.

This concept, now often referred to as lyonization, was not greeted with universal applause. Hans Gruneberg, a distinguished geneticist who made



**Mary Frances Lyon**

important contributions to understanding human disease through mouse genetics, simply did not believe the X inactivation hypothesis. Partly, it made no sense to have patches of inactivated cells (we now know that these are clones from early inactivation events), and partly he may have thought that Mary was too young and unknown to have such a fundamental insight. Gruneberg (a clever man, whose papers are worth re-reading) was 18 years older than Mary, well-known, and had been made a Fellow of the Royal Society (FRS) in 1956. He must have been a formidable opponent, but Mary stood her ground.

In the 1980s, Mary, who only ever worked with a small group, became involved with working out the human and mouse homology maps, alongside Harwell and Oxford scientists, including John Edward. This helped to lay the foundation for the Human Genome Sequencing Project 20 years or so later. Mary made major advances in understanding the inherited effects of radiation damage and seminal contributions to many areas of mouse genetics, including the t complex, embryo freezing with David Whittingham to preserve mouse stocks, and some early work with Richard Gardner in Cambridge on mouse preimplantation embryos to determine the timing of X inactivation.

Mary also edited Mouse News Letter (MNL) (1956–1970); in the postwar environment, UK and USA mouse geneticists in particular had close and long-standing

friendships and helped each other greatly. MNL was an informal publication mailed between them and other groups worldwide, providing each other with information about novel mutants and linkage groups—an important “bulletin board” in the pre-Internet world.

Mary won numerous awards and was elected an FRS in 1973 and a Foreign Associate of the US National Academy of Sciences in 1979 but was never one to push herself forward and did not receive the other honors she deserved. She was Head of Genetics at Harwell from 1962 to 1986, when she was glad to relinquish the increasing administrative burden to her successor, Bruce Cattanach. Under her stewardship, the Genetics Division had become a world-class center and was eventually established as the MRC Mammalian Genetics Unit, now directed by Steve Brown. Mary continued working productively at Harwell for many years following her official retirement in 1990. In 1998, Cambridge University held a degree ceremony at which Mary was awarded a full undergraduate degree. In 2004, the Mary Lyon Centre was opened at Harwell, in a fitting tribute to her contributions. Rather poignantly, the UK Genetics Society named a new medal after Mary last year, and though she knew of the medal, she did not live to meet its first awardee, Loeske Kruuk. Mary was diagnosed with Parkinson's disease a few years ago and died peacefully on Christmas Day, 2014.

This sums up the career, but not quite the woman. Mary was small and seemingly quiet. When one of us (E.M.C.F.) interviewed for a Ph.D position in 1983, Pete Glenister, Mary's long-time research assistant, warned “not to speak in the silences”—the silences were Mary's phenomenal mind at work, considering the data. When the other of us (J.P.) started work as a lab head at Harwell in 1978, Mary herself carried out the first set of ENU mutagenesis injections, just to help out. Mary had an extraordinary mind and memory and exceptional clarity. Her mild manner concealed an indomitable fighter on research matters where she felt strongly. She was a forward thinker and would have been thrilled to carry out experiments today using the panoply of research tools now available in mice. She also loved parties, never turning down an invitation. She was held in great affection by those who knew her, and she will be missed—an intellectual colossus, just over 5 feet tall.

**Elizabeth M.C. Fisher<sup>1,\*</sup>**  
**and Jo Peters<sup>2,\*</sup>**

<sup>1</sup>Department of Neurodegenerative Disease, Institute of Neurology, University College London, WC1N 3BG, UK

<sup>2</sup>Mammalian Genetics Unit, Medical Research Council, Harwell, Oxfordshire OX11 0RD, UK

\*Correspondence: [e.fisher@prion.ucl.ac.uk](mailto:e.fisher@prion.ucl.ac.uk) (J.P.)  
<http://dx.doi.org/10.1016/j.cell.2015.01.039>

# RNA: Jack of All Trades and Master of All

Julien Pompon<sup>1,\*</sup> and Mariano A. Garcia-Blanco<sup>1,2,\*</sup>

<sup>1</sup>Program of Emerging Infectious Diseases, Duke-NUS Graduate Medical School, Singapore 169857

<sup>2</sup>Department of Biochemistry and Molecular Biology, University of Texas Medical Branch, Galveston, TX 77555-0144, USA

\*Correspondence: julien.pompon@duke-nus.edu.sg (J.P.), maragarc@utmb.edu (M.A.G.-B.)

<http://dx.doi.org/10.1016/j.cell.2015.01.047>

**Noncoding RNAs have regulatory capabilities that evolution harnesses to fulfill diverse functions.** Lee et al. show that a noncoding RNA from Epstein-Barr virus recruits a host transcription factor to silence virus gene expression and propose that it does this through base-pairing with nascent viral transcripts.

As noted years ago by François Jacob, a broad set of processes that regulate gene expression appear to be the product of evolutionary tinkering (Jacob, 1977). For decades these mechanisms were thought to be exclusively protein-driven, but, as would be predicted by unfettered tinkering, many are now known to involve regulatory RNAs. These RNAs employ simple yet highly flexible modes of interaction with proteins and other nucleic acids to regulate every aspect of gene expression and function. In this issue of *Cell*, Lee et al. (2015) from the Steitz laboratory add a new trick in the repertoire of regulatory RNAs. The authors examine the function of an Epstein-Barr virus (EBV) noncoding RNA, *EBER2*, and, using capture hybridization analysis of RNA targets (CHART) (Lee et al., 2015 and references therein), find that *EBER2* localizes to the tandem terminal repeats (TRs) in the EBV genome, in the vicinity of where the PAX5 host transcription factor binds (Arvey et al., 2012). The authors go on to show that *EBER2* interacts with PAX5, albeit indirectly. Based on structure predictions, phylogenetic conservation in other related gamma herpesviruses, and experimental data, they also propose that *EBER2* forms an 18 bp hybrid with intronic TR sequences in viral *LMP2* nascent transcripts. This RNA-RNA interaction brings the *EBER2* associated PAX5 to the vicinity of its DNA binding site to enhance repression of LMP genes likely through chromatin remodeling (Figure 5 in Lee et al. 2015).

This provides a possible answer to the long open question regarding the function of the abundant *EBERs*. In that regard, several interesting questions are raised by the manuscript, does *EBER1* also

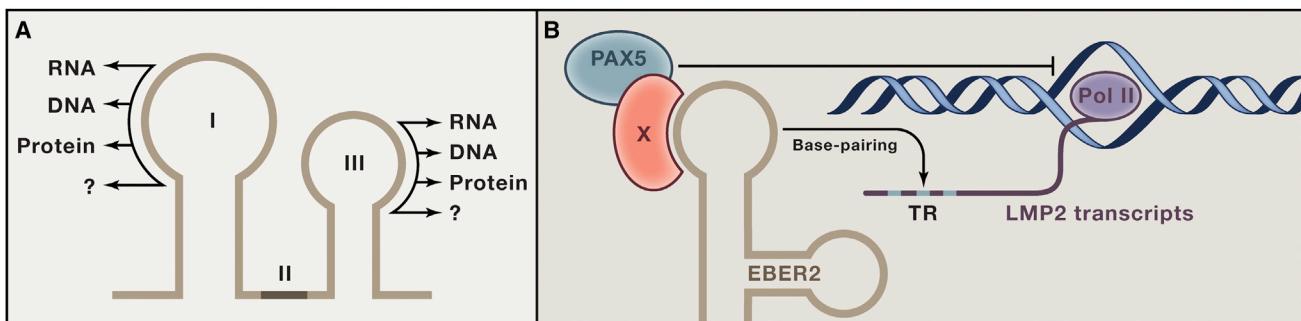
interact with PAX5? Indeed, careful inspection of Figure 2B in Lee et al. suggests that this may be the case. Could this explain the small effect of *EBER2* knockdown on PAX5 binding to the TR? As the authors themselves ponder—what about EBV strains deleted for *EBER2* (or both *EBERs*)? It is interesting to wonder whether the phenotypes observed with these strains (and there is controversy here) could be partially rescued by directly enhancing the PAX5 TR DNA interaction. These experiments would address the importance of *EBER*-mediated PAX5 recruitment for EBV replication and latency. As interesting as these questions are, the model of Lee et al. raises even more fascinating possibilities with general impact on RNA biology.

The model proposed in Figure 5 of Lee et al. represents a remarkable example of the versatile ability of RNAs to build complexes required for constitutive and regulated gene function. It also raises interesting questions. Can *EBER2* base pair with TR sequences in DNA, which would be accessible only when the region is transcribed? This scenario is not mutually exclusive with base-pairing to nascent RNAs, and one could imagine how the *EBER2* ribonucleoprotein would be handed from nascent RNA to DNA to bring PAX5 very close to its DNA binding site. Given the high density of nascent transcripts in many genomic regions, it is possible to imagine nascent RNAs as nets of binding sites that localize trans-activators near their eventual site of action. The ideas provoked by this manuscript add one more chapter to the rapidly evolving RNA story.

It is now clear that RNAs participate in almost every facet of the biology of cells

and viruses, and based on their function, RNAs have been categorized as protein-coding mRNAs or noncoding, which lack discernable open reading frames. Although this division is arbitrary and in many cases based on the absence of evidence, it has been widely used and serves as practical way to organize our rapidly changing understanding of RNA biology (Mercer et al., 2009). Excellent comprehensive reviews on noncoding RNAs (ncRNAs) and their many functions have been published (Mercer et al., 2009; Guttman and Rinn, 2012; Cech and Steitz, 2014).

Indeed, ncRNAs have many properties of adaptable regulators (Figure 1A): (1) RNAs, like DNAs, can “read” sequences by base-pairing and this ancient mode of nucleic acid-nucleic acid recognition provides very high specificity with minimal investment of genetic material. In contrast, proteins that “read” nucleic acid sequence generally do so by building complex binding domains (such as Puf proteins) (Wang et al., 2002). Additionally, RNAs have a proclivity to form structures that enhance base pairing and their 2' OH provides opportunities for hydrogen bonding. (2) RNAs interact with proteins using sequence, chemical modification of bases and sugars, and their secondary or tertiary structure. (3) RNAs, like proteins, are modular and can use domains or different surfaces within one domain to interact with other molecules (Guttman and Rinn, 2012). Furthermore, discrete interaction domains can be connected to form flexible modular scaffolds (Figure 1A). The modular nature of RNAs and the versatility of each module for diverse interactions dramatically expand the repertoire of regulatory RNAs and explain their exquisite specificity.



**Figure 1. RNA Modularity and Interaction Versatility**

(A) The schematic presents an RNA with two interaction modules (I and III) connected via a linker (II), which could be a hybrid linker in cases where I and III are in different molecules (e.g., CRISPRs). Each interaction module can interact with a diverse set of types of ligands. It is very likely that RNAs, like proteins, will be found to interact with every other type of macromolecule and small molecule present in cells (represented by the "?") (as already predicted by riboswitches and by the ability to select for binding to very different ligands *in vitro*).

(B) The example discovered in Lee et al. (2015) is presented in which an RNA molecule (*EBER2*) bridges between a protein (PAX5) and a second RNA (*LMP2*).

A modular RNA code, whereby discrete interaction domains can be combined into flexible modular scaffolds (I-III in Figure 1A) (Guttman and Rinn, 2012; Mercer et al., 2009), makes RNA a highly malleable substrate for evolutionary tinkering. This has been particularly apparent where rapid evolution is required as in host-pathogen interactions, such as the EBER2-PAX5 interaction. In fact, ncRNAs, encoded by both host and pathogen, play important roles in the control of the innate and acquired immune systems by altering every step of gene expression (Cech and Steitz, 2014). An excellent example of the modular evolution of RNA domains involved in host-pathogen interactions is provided by flaviviruses, such as dengue viruses, which cleave >90% of the genomes in infected cells to form a ncRNA derived from the 3' UTR. Elements in the 3' half of the ncRNA are conserved to serve in regulating translation of these viruses but elements in the 5' half, also known as the variable region, evolve rapidly to counter different components of host innate immunity (Bidet and Garcia-Blanco, 2014).

The resourcefulness of partner recognition by RNAs is exemplified by the EBER2: nascent TR:(X):PAX5 ribonucleoprotein (Figure 1B). EBER2 assembles with unknown factors (X) and hijacks PAX5. Additionally, EBER2 base-pairs with nascent RNAs to bring PAX5 in the vicinity of its DNA binding site—a new twist for *trans*-acting RNA. Small nuclear RNAs (snRNAs) are known to base-pair with nascent transcripts (pre-mRNAs) to mediate RNA splicing, as suggested by the Steitz group 35 years ago (Lerner et al., 1980), and HIV-1 Tat protein binds nascent TAR RNAs to recruit the cellular transcription factor P-TEFb to the lentiviral LTR (Wei et al., 1998). The modules described by Lee et al. are not new but the combination is—tinkering with any available part to build a new machine. We argue that the versatility of RNA makes it an exceptionally adept at sampling many forms and interactions that can assemble into a diverse array of machines, some of which will be selected. Whether or not RNA-based machines that recognize nascent transcripts are widely used is unclear. What is a foregone

conclusion, however, is that there are many RNA-based surprises ahead.

#### ACKNOWLEDGMENTS

The authors thank Micah Luftig for his comments on this manuscript.

#### REFERENCES

- Arvey, A., Tempera, I., Tsai, K., Chen, H.-S., Tikhmyanova, N., Klichinsky, M., Leslie, C., and Lieberman, P.M. (2012). *Cell Host Microbe* 12, 233–245.
- Bidet, K., and Garcia-Blanco, M.A. (2014). *Biochem. J.* 462, 215–230.
- Cech, T.R., and Steitz, J.A. (2014). *Cell* 157, 77–94.
- Guttman, M., and Rinn, J.L. (2012). *Nature* 482, 339–346.
- Jacob, F. (1977). *Science* 196, 1161–1166.
- Lee, N., Moss, W.N., Yario, T.A., and Steitz, J.A. (2015). *Cell* 160, this issue, 607–618.
- Lerner, M.R., Boyle, J.A., Mount, S.M., Wolin, S.L., and Steitz, J.A. (1980). *Nature* 283, 220–224.
- Mercer, T.R., Dinger, M.E., and Mattick, J.S. (2009). *Nat. Rev. Genet.* 10, 155–159.
- Wang, X., McLachlan, J., Zamore, P.D., and Hall, T.M.T. (2002). *Cell* 110, 501–512.
- Wei, P., Garber, M.E., Fang, S.M., Fischer, W.H., and Jones, K.A. (1998). *Cell* 92, 451–462.

# Cell Migration, Freshly Squeezed

Matthew D. Welch<sup>1,\*</sup>

<sup>1</sup>Department of Molecular and Cell Biology, University of California, Berkeley, Berkeley, CA 94720, USA

\*Correspondence: [welch@berkeley.edu](mailto:welch@berkeley.edu)

<http://dx.doi.org/10.1016/j.cell.2015.01.053>

Migrating cells exhibit distinct motility modes and can switch between modes based on chemical or physical cues. Liu et al. and Ruprecht et al. now describe how confinement and contractility influence motility mode plasticity and instigate a mode termed stable bleb migration in embryonic and tumor cells.

Crawling cell migration arose early in the evolution of eukaryotes and is exhibited by diverse eukaryotic lineages (Fritz-Laylin et al., 2010). In multicellular animals, cell migration is crucial in development, tissue maintenance, and immunity. Cell migration gone awry can also contribute to disease, for example, during cancer invasion and metastasis (Friedl and Alexander, 2011). The migration behaviors of animal cells are heterogeneous and differ depending on cell type, developmental stage, local environment, and disease state. In this issue of *Cell*, studies from the Piel group (Liu et al., 2015) and Heisenberg group (Ruprecht et al., 2015) uncover how the plasticity of cell migration behaviors is impacted by the strength of adhesion, physical confinement (squeezing between two surfaces), contractility, and chemical cues.

To make sense of the heterogeneity of migration behaviors, the migration modes of single cells have been divided into two broad categories—mesenchymal and amoeboid—based on cell morphology, mechanism of force generation, cytoskeleton organization, and characteristics of the cell-substrate interaction (Figure 1) (Friedl and Wolf, 2010). The mesenchymal mode, exemplified by fibroblasts, is typified by slower velocity, irregular shape, strong cell-substrate adhesion, prominent actin stress fibers, and actin-rich leading edge structures, including lamellipodia and filopodia (Gardel et al., 2010). The amoeboid category is characterized by a faster velocity, rounder shape, weaker cell-substrate adhesion, and the absence of stress fibers (Lämmermann and Sixt, 2009). Distinct amoeboid motility modes, however, employ different leading edge structures. Immune cells typically form actin-rich pseudopods at

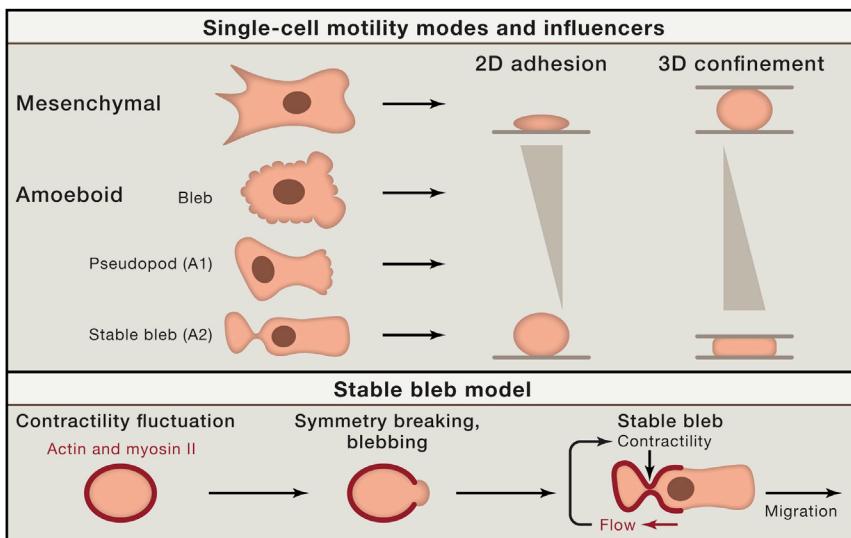
their leading edge (Figure 1, pseudopod mode), for which protrusion is driven by local actin assembly. Other cells—zebrafish primordial germ cells, for example—instead form actin-deficient blebs at their leading edge, for which protrusion is driven in part by contraction at the rear by actin and myosin II (Figure 1, bleb mode) (Paluch and Raz, 2013). There is plasticity between motility modes, and cells can switch between modes, depending on their environment (Friedl and Wolf, 2010).

To investigate the effects of adhesion strength and physical confinement on cell migration plasticity, Liu et al. (2015) systematically vary the strength of adhesion and degree of confinement of a fibroblast cell line that normally migrates in the mesenchymal mode. They observe that, under conditions of low 2D adhesion and high 3D confinement (Figure 1), fibroblast cells switch from mesenchymal to one of two higher-velocity amoeboid motility modes, a so-called mesenchymal-to-amoeboid transition (MAT). Cells migrating in the A1 mode have a round cell body and small leading edge resembling a pseudopod or lamellipod. Unexpectedly, cells migrating in the A2 mode have an unusually long and ellipsoid cell body and leading edge. Examination of numerous other cell lines shows that many favor one mode or the other, but some exhibit both modes. Notably, transformed and tumor cells, as well as leukocytes, prefer the A2 mode. These findings suggest that individual cell lines exhibit migration mode plasticity and that adhesion strength and confinement are key parameters in controlling migration behavior.

At the molecular level, Liu et al. (2015) find that the ability of tumor cells to switch to the A2 mode depends on diminished

activity of proteins that contribute to cell-substrate adhesion. Moreover, conditions of high contractility generated by myosin II favor the A2 mode, whereas conditions of low contractility favor the A1 mode. The driving force for the A2 mode is linked to global retrograde flow of actin and myosin II in the central region of the cell, whereas the A1 mode is linked to retrograde flow only in the leading edge. In A2 cells, the result of more global cortical flow of actin and myosin II is the accumulation of these proteins at the cell rear and the formation of a stable bleb at the leading edge that is largely devoid of actin. Mathematical modeling also predicts the appearance of this stable bleb in a regime of high contractility. In a previous study, cells migrating with a stable bleb-like protrusion dependent on myosin II function were observed in *Dictyostelium discoideum* treated with quinine, although the physiological significance remained unclear (Yoshida and Inouye, 2001). The Liu et al. (2015) study advances this earlier work by examining physiologically relevant environmental factors that influence migration plasticity and stable bleb migration and by providing a mechanistic description of motility.

In a companion study, Ruprecht et al. (2015) examine the migration modes of primary germ layer progenitor cells isolated from zebrafish embryos. They find that cells plated on a 2D substrate undergo amoeboid blebbing or mesenchymal motility, depending on the germ layer origin of the cells and the extracellular matrix composition of the substrate. However, if cells of various origins are treated with serum or the serum component lysophosphatidic acid (LPA), they switch to a rapid amoeboid motility



**Figure 1. Motility Modes, Influencing Parameters, and the Initiation and Maintenance of Stable Bleb Motility**

Motility modes include mesenchymal motility, or various forms of amoeboid motility characterized by blebs, pseudopods, or stable blebs. These are influenced by environmental factors, including the strength of adhesion to the substrate, or the extent of physical confinement and contractility. The formation of a stable bleb is theorized to involve fluctuations in cortical contractility, followed by symmetry breaking and the formation of a bleb. Cortical flow of actin and myosin II toward the cell rear then enhances contractility in this location, stabilizing the formation of a single bleb at the front and generating force that drives migration.

mode characterized by the formation of a single stable bleb, similar to the A2 mode described by Liu et al. (2015). Spatial confinement of zebrafish cells in the absence of serum also induces stable bleb formation and migration. Furthermore, for both LPA-induced and confinement-induced stable bleb migration, myosin II accumulates at the rear of the cell, and myosin-II-driven contractile activity is required. These results, together with those of Liu et al. (2015), indicate that stable bleb motility is observed in primary cells and cell lines isolated from various animal species.

To understand how cortical contractility contributes to stable bleb motility, Ruprecht et al. (2015) develop a mathematical description of this process. Their model predicts that local fluctuations in cortical contractility at the cell periphery are amplified by external cues or physical confinement, causing symmetry breaking and initial polarization (Figure 1). Cell polarization is then enhanced and stabilized by positive feedback between cortical flow of actin and myosin II toward the

cell rear and the formation of a cortical contractility gradient that reinforces flow. In support of this model, they observe coupling of cortical flow of actin and myosin to stable bleb migration. Thus, the theoretical model and experimental evidence suggest that contractility itself is sufficient to initiate stable bleb motility.

They then examine the incidence of stable bleb motility in zebrafish embryos in response to contractility. Cells expressing a constitutively active variant of the Rho family G protein RhoA (to enhance cortical contractility) exhibit similar shape, migration behavior, and cortical actin and myosin II flows as stable bleb cells in vitro. Moreover, when transferred into cell culture, these cells exhibit characteristics indistinguishable from primary germ layer cells undergoing stable bleb motility. Notably, at wounding sites in embryos, which exhibit high levels of contractility, cells not expressing exogenous RhoA also assume a stable bleb motility mode and migrate from the wounding site with the rapid speed, directional persistence, and characteristics of

stable bleb cells in vitro. This confirms that stable bleb motility initiates at regions of high contractility in a developing embryo.

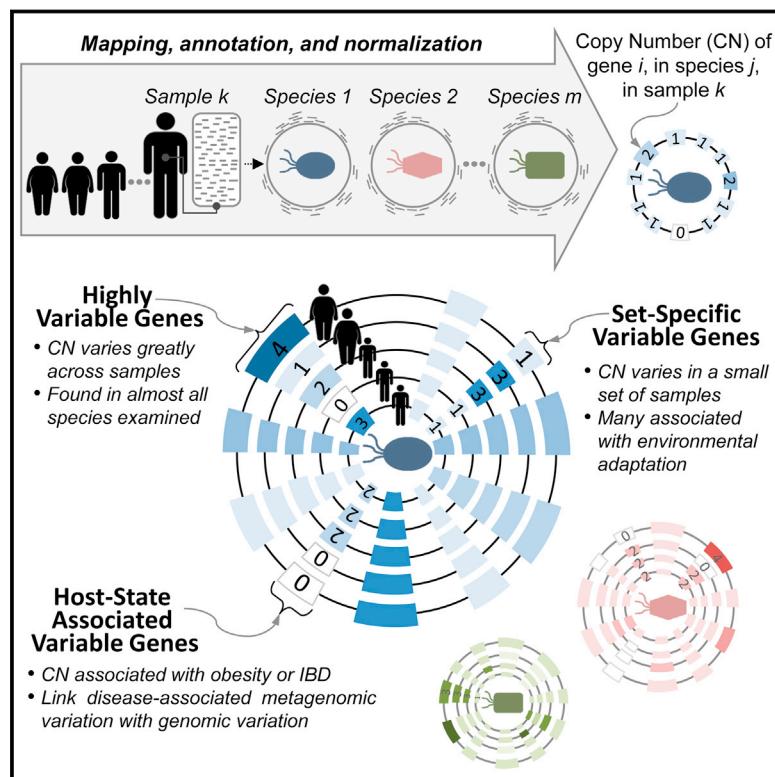
The findings of Liu et al. (2015) and Ruprecht et al. (2015) for animal cells, along with the previous study of *Dictyostelium* amoeba (Yoshida and Inouye, 2001), suggest that stable bleb motility is a fundamental motility mode of eukaryotic cells and that it operates in a variety of contexts. These may include embryonic development, where it is speculated that this rapid and directionally persistent motility enables extrusion of cells from regions of high contractility and mediates long-range cell interactions. Moreover, stable bleb motility is prevalent in transformed and tumor cells subjected to confinement, suggesting that it may be a fundamental property related to invasion and metastasis. Future work will establish the detailed mechanisms of force generation during stable bleb motility, the contribution of this process to development and tissue maintenance, and the impact of this form of migration on cancer and other diseases.

## REFERENCES

- Friedl, P., and Alexander, S. (2011). *Cell* 147, 992–1009.
- Friedl, P., and Wolf, K. (2010). *J. Cell Biol.* 188, 11–19.
- Fritz-Laylin, L.K., Prochnik, S.E., Ginger, M.L., Dacks, J.B., Carpenter, M.L., Field, M.C., Kuo, A., Paredez, A., Chapman, J., Pham, J., et al. (2010). *Cell* 140, 631–642.
- Gardel, M.L., Schneider, I.C., Aratyn-Schaus, Y., and Waterman, C.M. (2010). *Annu. Rev. Cell Dev. Biol.* 26, 315–333.
- Lämmermann, T., and Sixt, M. (2009). *Curr. Opin. Cell Biol.* 21, 636–644.
- Liu, Y.-J., Le Berre, M., Lautenschlaeger, F., Maiuri, P., Callan-Jones, A., Heuzé, M., Takaki, T., Voituriez, R., and Piel, M. (2015). *Cell* 160, this issue, 659–672.
- Paluch, E.K., and Raz, E. (2013). *Curr. Opin. Cell Biol.* 25, 582–590.
- Ruprecht, V., Wieser, S., Callan-Jones, A., Smutny, M., Morita, H., Sako, K., Barone, V., Ritsch-Marte, M., Sixt, M., Voituriez, R., and Heisenberg, C.-P. (2015). *Cell* 160, this issue, 673–685.
- Yoshida, K., and Inouye, K. (2001). *J. Cell Sci.* 114, 2155–2165.

# Extensive Strain-Level Copy-Number Variation across Human Gut Microbiome Species

## Graphical Abstract



## Authors

Sharon Greenblum, Rogan Carr, Elhanan Borenstein

## Correspondence

elbo@uw.edu

## In Brief

Extensive strain-level variation is detected in the human gut microbiome, with differences in gene copy-number impacting specific adaptive functions and linked to obesity and inflammatory bowel disease.

## Highlights

- A metagenomic data analysis pipeline allows strain-level gene copy-number inference
- Copy-number variation (CNV) is widespread across many prevalent human gut species
- CNV involves mostly environment-related functions and is associated with disease
- Strain-level population structure reveals known and uncharacterized strains



# Extensive Strain-Level Copy-Number Variation across Human Gut Microbiome Species

Sharon Greenblum,<sup>1</sup> Rogan Carr,<sup>1</sup> and Elhanan Borenstein<sup>1,2,3,\*</sup>

<sup>1</sup>Department of Genome Sciences, University of Washington, Seattle, WA 98195, USA

<sup>2</sup>Department of Computer Science and Engineering, University of Washington, Seattle, WA 98195, USA

<sup>3</sup>Santa Fe Institute, Santa Fe, NM 87501, USA

\*Correspondence: elbo@uw.edu

<http://dx.doi.org/10.1016/j.cell.2014.12.038>

## SUMMARY

Within each bacterial species, different strains may vary in the set of genes they encode or in the copy number of these genes. Yet, taxonomic characterization of the human microbiota is often limited to the species level or to previously sequenced strains, and accordingly, the prevalence of intra-species variation, its functional role, and its relation to host health remain unclear. Here, we present a comprehensive large-scale analysis of intra-species copy-number variation in the gut microbiome, introducing a rigorous computational pipeline for detecting such variation directly from shotgun metagenomic data. We uncover a large set of variable genes in numerous species and demonstrate that this variation has significant functional and clinically relevant implications. We additionally infer intra-species compositional profiles, identifying population structure shifts and the presence of yet uncharacterized variants. Our results highlight the complex relationship between microbiome composition and functional capacity, linking metagenome-level compositional shifts to strain-level variation.

## INTRODUCTION

The human gut microbiome plays an important role in host metabolism, immunity, and drug response and has a tremendous impact on our health (Iida et al., 2013; Kinross et al., 2011; Vijay-Kumar et al., 2010). Numerous comparative studies aiming to characterize the contribution of the microbiome to human health have already demonstrated marked shifts in the relative abundance of various species, genera, or phyla in various disease states (Frank et al., 2007; Hoffman et al., 2014; Larsen et al., 2010; Turnbaugh et al., 2009). Clearly, however, each microbial species represents many different strains that may encode considerably different sets of genes and a different number of copies of each gene (reflecting, for example, gene deletions and duplication events). Such intra-species variation endows each strain with potentially distinct functional capacities. Studies of individual isolates of cultured species have indicated, for example, that strains often differ in virulence (Gill et al., 2005;

Salama et al., 2000; Solheim et al., 2009), motility (Zunino et al., 1994), nutrient utilization (Siezen et al., 2010), and drug resistance (Gill et al., 2005). Accordingly, the true functional potential of a microbiome cannot be inferred from species composition alone, and species-level comparative analyses may fail to capture important functional differences across samples. Recent efforts to catalog the relative abundance of known strains in human microbiome samples (Kraal et al., 2014) may recover some of these differences but are limited to sequenced reference genomes and are not able to identify novel, yet-to-be-sequenced variation. Gene-centric shotgun metagenomic studies, on the other hand, may identify genes or pathways that are differentially abundant across samples but cannot necessarily attribute these shifts to specific species or strains. Specifically, it is often unclear how much of the observed variation in gene composition is due to variation in the abundances of species and how much is contributed by intra-species variation. Indeed, conflicting results have been reported, with trends identified among species profiles that are often poorly translated to gene profiles and vice versa (Muegge et al., 2011; Turnbaugh et al., 2009). It is therefore not yet clear how prevalent gene-level intra-species variation is in the human gut, whether such variation is adaptive and affects specific functions, and how much of this variation has already been captured by reference genomes.

Some evidence already suggests that variation among strains is common in the human gut. Several studies have focused specifically on nucleotide-level variation, assessing, for example, the prevalence and stability of single-nucleotide polymorphisms across numerous metagenomes (Schloissnig et al., 2013) or the level of sequence diversity across multiple near-complete genomes from two bacterial species variants obtained by single-cell sequencing (Fitzsimons et al., 2013). Other studies have taken steps to associate sequence variation with gene-level differences, identifying, for example, areas of variable coverage and the coordinated loss of genes from specific gene families within the *Streptococcus mitis* B6 genome (Human Microbiome Project Consortium, 2012) or a diverse array of strain-specific adhesion-like protein genes across cultured strains of *Methanobrevibacter smithii* (Hansen et al., 2011). Additional studies have used extensive manual genomic reconstruction to track strain-resolved shifts over time in *Actinomycetaceae* in the relatively low-complexity premature infant gut microbiome (Brown et al., 2013); to detect differences related to antibiotic resistance, transport, and biofilm formation among three strains of *Staphylococcus epidermidis* (Sharon et al., 2013); or to identify

the variable presence of genes involved in transport, motility, carbohydrate metabolism, and virulence in two distinct strains of *Citrobacter* (Morowitz et al., 2011). These gene-level studies, however, mostly report small-scale or anecdotal results, focusing on one or a small number of species and often on specific gene families. A high-throughput, comprehensive analysis of gene-level variation across a large array of species in the human gut is therefore needed to more fully appreciate the extent and functional implications of strain variation in this complex microbiome.

To address this challenge, here we establish a rigorous and robust pipeline to estimate the copy number of each gene in a large set of prevalent gut microbial species in a given sample directly from metagenomic shotgun data and, furthermore, to detect copy-number variation across samples. We carefully calibrate this pipeline to confirm that it can successfully estimate the copy number of individual genes in individual species on a large scale. Applying this pipeline to 109 metagenomic samples from a recent study of the gut microbiomes of healthy, obese, and inflammatory bowel disease (IBD)-afflicted individuals, we estimate the copy number of more than 4,000 gene groups across 70 species in each of these samples and demonstrate the presence of widespread copy-number variation within many genes in many species. We find that specific functions are especially prone to copy-number variation, including functions relevant to a community lifestyle and adaptation to the gut environment, and further detect associations between strain variation and host phenotype. Finally, we demonstrate that these copy-number estimates can be used both to model the composition of known strains within each sample and to offer insight into complex population structures, suggesting the presence of yet uncharacterized species variants.

## RESULTS

### A Pipeline for Calculating Genomic Copy-Number Estimates in Metagenomic Samples

We developed a pipeline to confidently detect variation in gene content and gene copy number in a large set of prevalent human gut microbes directly from metagenomic data (Figure 1 and Experimental Procedures). Briefly, this pipeline works as follows. Shotgun metagenomic short reads were first aligned to a set of reference genomes representing dominant and prevalent gut microbiome strains. To account for the potentially multiple genomes available for each species in this reference database, genomes were grouped into clusters using a previously introduced sequence similarity-based method (Schloissnig et al., 2013). These clusters represent approximate species-level groups, though in some cases may not reflect classical taxonomic divisions. We used extensive simulations to carefully select alignment parameters and confirmed that, with these parameters, reads mapped to the correct region and correct genome cluster, whereas reads from genome clusters not represented in our reference database remained unmapped (Figure 2A; Figure S1; Extended Experimental Procedures). In parallel, gene coding regions from all reference genomes were annotated with KEGG orthology groups (KOs). Reference genomes and KOs with low confidence mapping were identified and excluded (Figure S2; Extended Experimental Procedures). For each sample, coverage

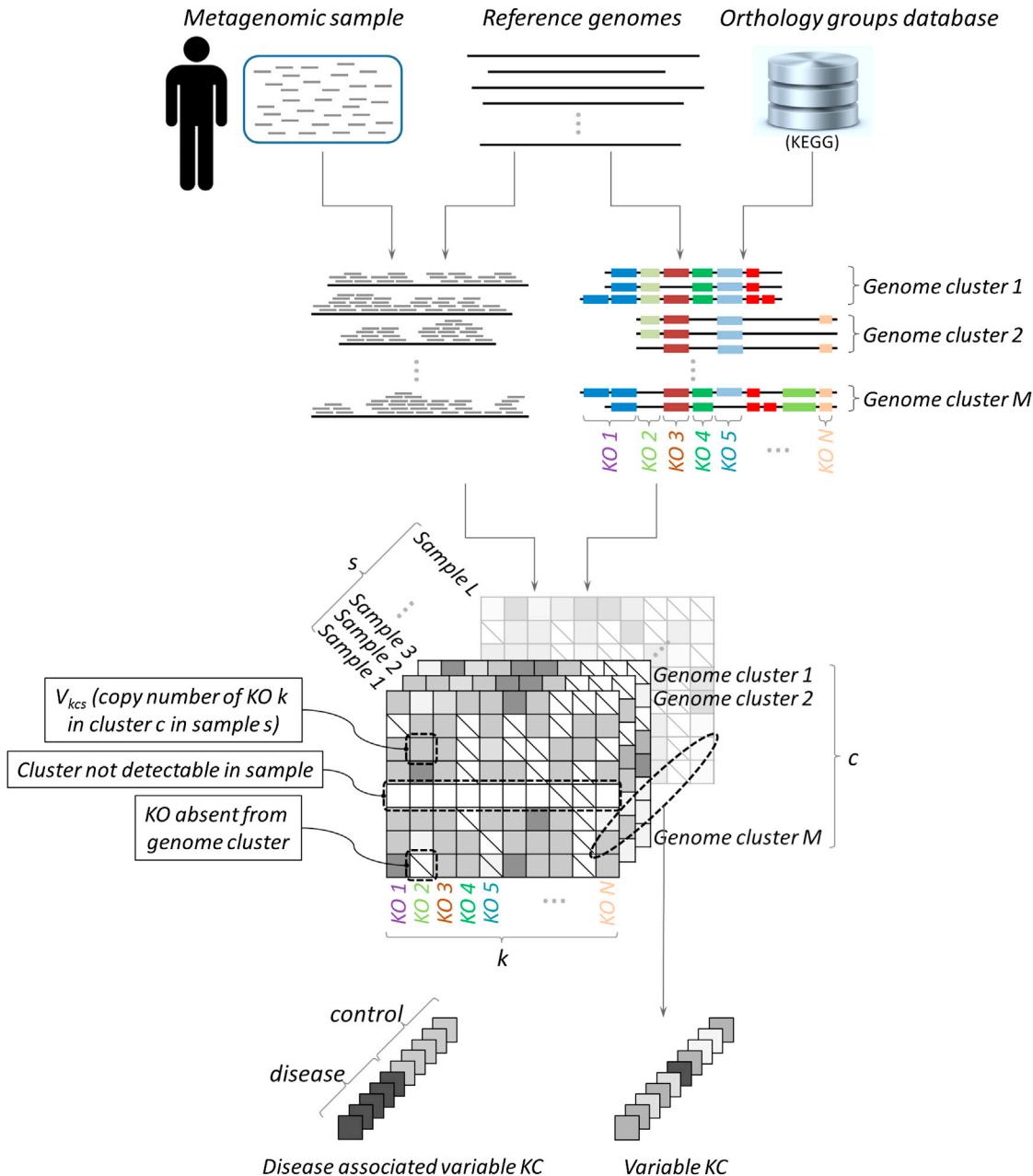
across each KO-annotated region in each reference genome was calculated, and coverage values across regions corresponding to the same KO in the same genome cluster were summed. We then used the average coverage of 13 single copy marker genes, carefully selected for their universality, mapping accuracy, and coverage consistency (Figure S3; Extended Experimental Procedures), to convert the calculated coverage of each KO in each cluster to a copy-number estimate (Experimental Procedures). Overall, this process estimated the copy number,  $V_{kcs}$ , of each KO  $k$ , in each genome cluster  $c$ , detected in each sample  $s$  (Figure 1). Notably, copy-number estimates represent an average across the various genomes associated with each cluster that are present in the sample and across the potentially multiple genes associated with each KO. We further performed an analysis of an extensive synthetic dataset to confirm that this scheme accurately recovers species abundances and copy-number values (Figures S4A and S4B; Extended Experimental Procedures).

We applied this pipeline to a dataset of 109 previously collected gut metagenomic samples from a Danish/Spanish cohort (Qin et al., 2010), mapping in total >2.45 billion 75 bp reads to 235 reference genomes grouped into 96 genome clusters (Table S1; Extended Experimental Procedures). The average coverage across the 13 marker genes (a proxy for the abundance of each cluster in each sample) varied considerably across clusters and between samples (Figures 2B and 2C). To limit any downstream analysis to high-confidence copy-number estimates, we therefore considered only genome clusters with sufficient coverage in a sample (which we term “detectable” clusters; Experimental Procedures). We identified a total of 70 clusters that were detectable in at least one sample, with an average of 16 detectable clusters in each sample (Table S2). Overall, this analysis assigned copy number values to ~1.5 million KO-cluster-sample triplets, estimating the copy number of thousands of KOs across a large array of genome clusters in >100 samples (Table S3).

This dataset of copy-number estimates provides a first large-scale account of gene-level strain variation among organisms common to the human gut. Below, we mine this dataset to explore neutral and adaptive variation in this highly complex ecosystem in a manner that goes beyond species-level comparative analysis. Importantly, this dataset and the pipeline described above can serve as a valuable resource for future studies of compositional shifts in the human microbiome and in other environments, linking metagenome-level differences in gene abundance to genome-level variation.

### Identifying Genes with Highly Variable and with Set-Specific Variable Copy Number

Given the copy-number estimates obtained above, we set out to identify specific KOs in specific clusters (KO-cluster pairs, or KCs) whose copy number varied across samples. Notably, to detect variation, we compared the copy number of each KC across different samples rather than comparing the estimated copy number in any given sample to the copy number in a reference genome, avoiding spurious variation predictions that may result from annotation errors or bias in the set of reference genomes. Clearly, many clusters can be detected in only a few samples. To confidently detect copy-number variation, we



**Figure 1. Schematic of Analysis Pipeline**

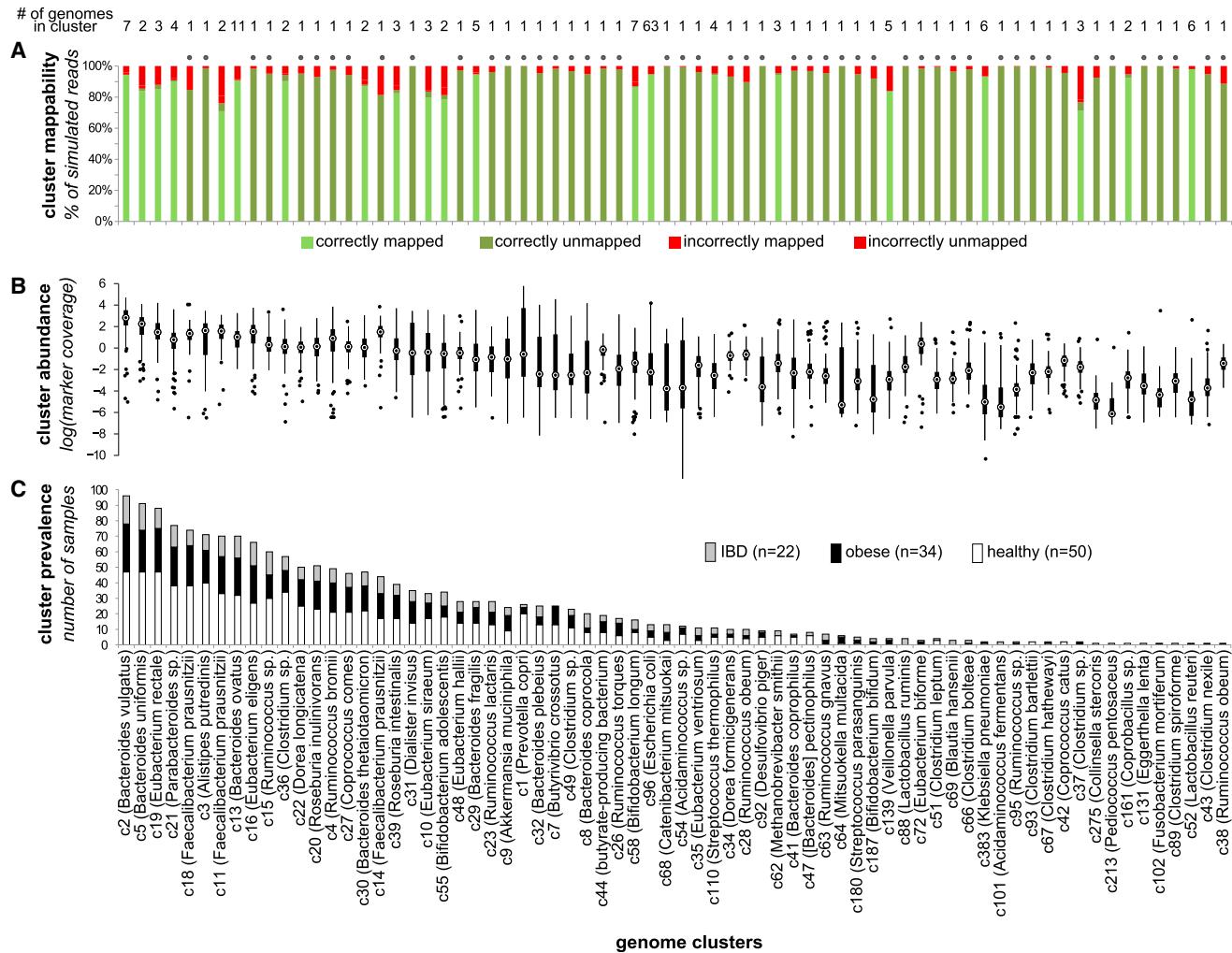
Reads from metagenomic samples were mapped to KEGG-annotated reference genomes, grouped into species-level genome clusters. The total coverage of each KO (KEGG orthology group),  $k$ , in each genome cluster,  $c$ , in each sample,  $s$ , was normalized by cluster abundance to calculate gene copy number  $V_{kcs}$ . KCs (specific KOs in specific genome clusters) whose copy number varied significantly across samples were detected, as well as those whose copy number was associated with host state (obesity, IBD).

See also Figure S3 and Table S3.

therefore only considered the 40 clusters that were detectable in at least 10 samples.

We first set out to identify KCs that exhibit extreme and prevalent variation across samples. Specifically, we calculated the level of inter-sample variation in the copy number of each KC and defined as *highly variable* those KCs whose variation was at least two standard deviations greater than the average

variation of all KCs (Experimental Procedures). We used both cross-validation analysis and synthetic samples to confirm the robustness and accuracy of this approach (Extended Experimental Procedures; Figure S4C). In total, this analysis detected 735 highly variable KCs spanning 261 KOs across 38 genome clusters (Figure 3; Table S4). The number of highly variable KCs in each cluster varied greatly, reaching up to 47 KCs in



## Figure 2. Genome Cluster Statistics

The mappability, abundance, and prevalence of each genome cluster (representing a species-level group) are shown in three vertically aligned plots. Clusters are sorted by their prevalence across samples.

(A) Cluster mappability, as determined by a large-scale simulation assay measuring the accuracy of mapping reads extracted from the cluster's genomes to a database in which the genome of origin was removed. In this simulation, reads from clusters represented in the reference database by a single genome (marked with a dot above the column) are expected to remain unmapped.

(B) The distribution of each cluster's abundance across samples, as determined by the average coverage of 13 single-copy marker genes.

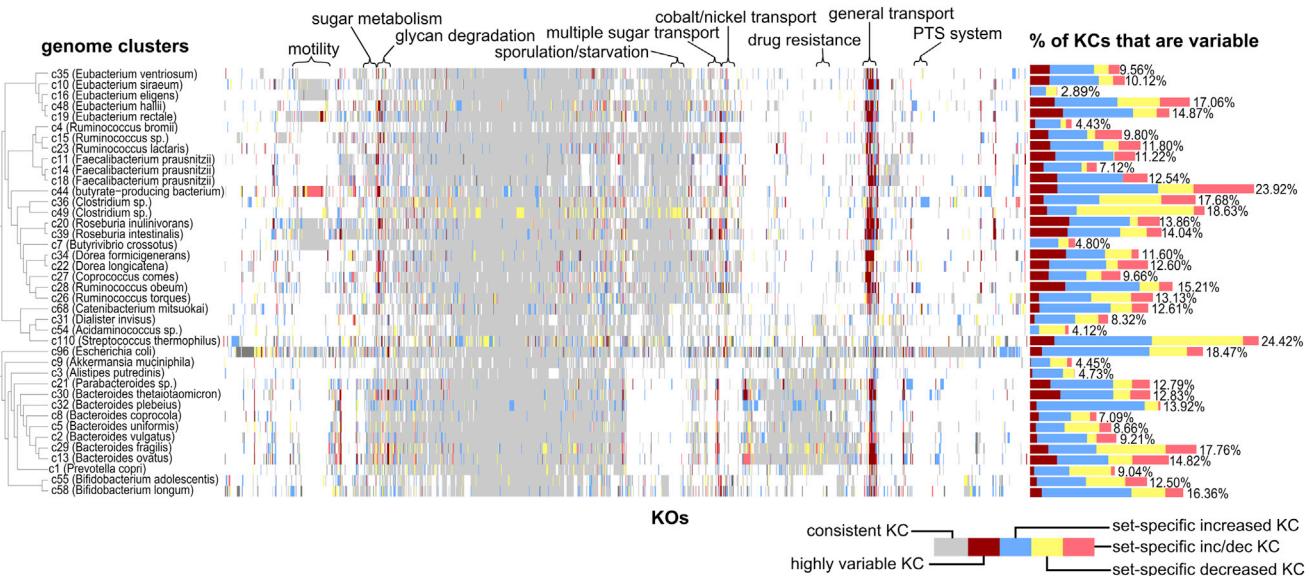
(C) Cluster prevalence (the number of samples in which the cluster was “detectable”) within each host group, shown as a stacked bar plot.

See also Figures S1 and S2 and Tables S1 and S2.

the *Roseburia intestinalis* cluster (representing 4.05% of the KCs in this cluster), with an average of 1.79% of the KCs in each cluster (Table S5). We found no apparent relationship between the amount of variation observed in a cluster and the number of reference genomes in the cluster or the prevalence of the cluster across samples, but we did observe a tendency toward high variation in species from the *Firmicutes* phylum compared to other species (*t* test,  $p < 0.05$ ; see also Figure 3). Although the majority of highly variable KOs (57.1%) were variable in just one cluster, certain KOs were variable across many clusters, with some KOs variable in ten or more different clusters.

The analysis above focused on KCs that exhibit extreme variation and on KCs that vary greatly across *many* different samples.

Variation within other genes, however, may be more subtle and may reflect, for example, adaptive variation that can be observed in only a small set of samples. We therefore set out to additionally identify set-specific variable KCs, wherein the copy number of a given KC was relatively constant across most samples but deviated significantly in a small subset of the samples ([Experimental Procedures](#)). In this analysis, we further distinguished cases in which a KC exhibited a consistently high copy number in this subset of samples compared to all other samples (set-specific increased copy number) from cases in which a KC exhibited a consistently low copy number in this subset of samples (set-specific decreased copy number) or in which it exhibited increased copy number in one subset and decreased in another. As



**Figure 3. A Map of Variable KCs**

A matrix map representing the status of variable KOs (x axis) in each genome cluster (y axis). Colored bars represent variable KCs (highly variable KCs vary widely in copy number across all samples, whereas set-specific variable KCs are increased and/or decreased in copy number in only a small subset of the samples), while light gray bars indicate KCs with consistent copy number across samples, and KOs not present in a genome cluster are left white. Genome clusters are ordered by phylogeny, and KOs are ordered by hierarchical clustering. The bar chart to the right of the map represents the fraction of KOs in each cluster identified as variable. Above the map, certain groups of functionally related KOs are highlighted. The 314 KOs uniquely variable in the *E. coli* cluster (the majority of which have only been annotated in *E. coli*-like genomes) were excluded due to space constraints.

See also Figure S4 and Tables S4–S6.

expected, we found that set-specific variable KCs were much more common than highly variable KCs. In total, our analysis detected 5,004 set-specific variable KCs covering 1,859 KOs across the 40 genome clusters examined (Figure 3; Table S4). In general, we observed more cases of set-specific increased copy number than of set-specific decreased copy number, but this ratio shifted markedly across clusters, and in certain clusters (i.e., *Clostridium* sp., *Streptococcus thermophilus*) mostly set-specific decreased KCs were observed.

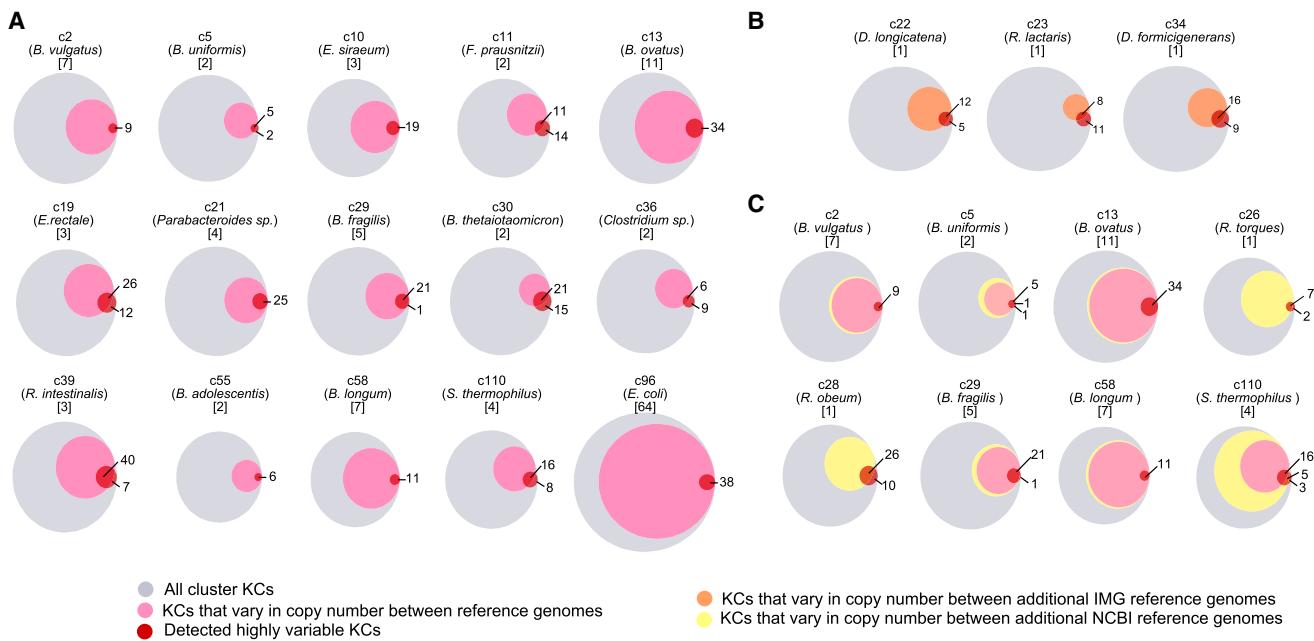
### Detected Variation Captures Both Known and Novel Strain Variation

As validation of our pipeline and results, we compared the set of highly variable KCs obtained for each cluster to known variation among the cluster's sequenced reference genomes. Clearly, the reference genomes in our database do not capture the full extent of intra-species variation in the gut microbiome. Similarly, our samples likely do not include much of the variation present in our reference genomes, as many of these reference genomes represent strains isolated from clinically distinct individuals, phenotypically diverse cohorts, or non-gut samples. Accordingly, a large number of genes that vary in copy number across reference genomes may still exhibit consistent copy number across the gut samples analyzed above. Yet, the set of detected highly variable genes, which aims to include genes that vary frequently in their copy number across genomes, is likely to capture many instances of known variation in gene content among available reference genomes. Indeed, considering the 15 multiple-genome clusters in our database, a striking 81% of the detected highly variable

KCs also vary in copy number across reference genomes (Figure 4). Moreover, in seven of these clusters, all highly variable KCs also vary in copy number across reference genomes. Notably, six of these clusters contain at least three genomes, whereas the majority of the other clusters contain only two, suggesting that more sequenced strains may be needed to fully capture the variation associated with these clusters (and more importantly, with clusters for which only a single genome was available). Importantly, we demonstrated that a similar overlap can be observed when comparing predicted variation to known variation among a large collection of genomes not included in our database, confirming that this overlap is not an artifact of the specific reference genomes used in our analysis (Figures 4B and 4C; Extended Experimental Procedures). Comparison of set-specific variable KCs to known variation across reference genomes again confirmed that the variation detected greatly overlapped with known variation observed across sequenced strains (Figure S5). Notably, however, set-specific variable KCs also included many instances of novel variation, suggesting that the set of reference genomes currently available does not capture the full extent of copy-number variation in the gut. Comparison of detected set-specific variation to variation observed across two manually assembled *Citrobacter* strains further revealed significant overlap (Extended Experimental Procedures).

### Functions Associated with Variable Genes

We examined whether the detected copy-number variation was associated with specific functions in each genome cluster. We first used enrichment analysis to identify functions that were



**Figure 4. Comparison of Highly Variable KCs to Known Variation among Reference Genomes**

(A) In each Venn diagram, the gray circle represents the set of all KCs in a given genome cluster, the pink circle represents the fraction of those KCs exhibiting copy-number variation across the cluster's reference genomes, and the red circle represents the set of KCs detected as highly variable. Overlap of the pink and red circles indicates correspondence between known and detected variation. Each diagram is labeled with the cluster ID, representative species name, and number of reference genomes.

(B and C) Additional variation in reference genomes that were not used as mapping targets is represented by either an orange circle (additional reference genomes from IMG) or a yellow circle (additional reference genomes from NCBI), compared to variation in included reference genomes (pink) and detected highly variable KCs (red).

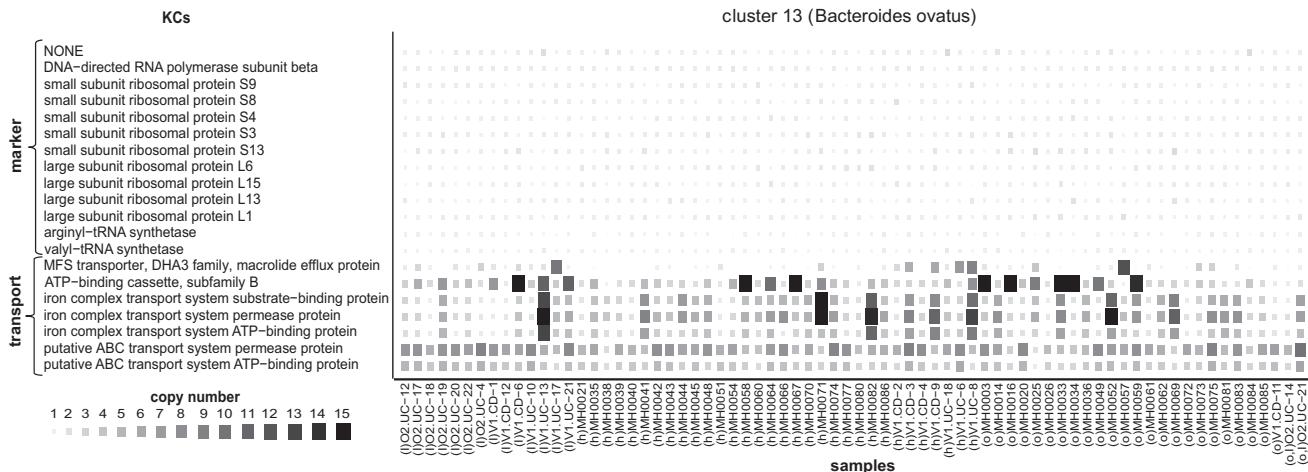
See also Figure S5.

over-represented among the set of highly variable KCs in each cluster. We found that transport-related functions were overwhelmingly prone to high copy-number variation (Table S6). Specifically, ten of the genome clusters analyzed were enriched for variation in KCs associated with transport annotations, including the general BRITE term "Transporter," as well as more specific modules related to either sugar or iron complex transport. For example, within the *Bacteroides ovatus* cluster, seven of the cluster's 66 transport-associated KCs were highly variable (Figure 5), including all three KCs (K02013, K02015, K02016) involved in a specific iron complex transport system module (M00240). Interestingly, significant variation in sugar transport functions was only found among clusters in the phyla Firmicutes and Actinobacteria, whereas Bacteroidetes clusters were uniquely associated with variation in the iron complex transport system (see Table S6). Studies of cultured organisms from various environments and experimental evolution assays have suggested that loss, amplification, and acquisition of transport functions constitute a primary adaptive mechanism (Gevers et al., 2004; Heikkinen et al., 2007; Lee and Marx, 2012; Sonti and Roth, 1989); here, we show that this flexibility in the copy number of transport genes likely extends to a considerable proportion of prevalent gut species and that, within this general class, specific transport genes may facilitate adaptation to the gut environment.

We additionally found that motility-related KCs were highly variable in the *Eubacterium rectale* genome cluster. Specifically,

in this cluster, 7 of the 38 highly variable KCs were bacterial motility proteins, of which four were structural flagellar components, two were involved in chemotaxis, and one was essential for twitching motility (Han et al., 2008). Motility proteins, and especially flagellar proteins, are widely associated with virulence and immunostimulation, and the gain or loss of flagellar components is believed to be an important adaptive mechanism (Borziak et al., 2013; Heikkinen et al., 2007; Al Mamun et al., 1997). Moreover, variation in these seven KCs was highly consistent within samples; most samples contained either detectable copies of all seven KCs or no (or low number of) copies of all of these KCs (Figure S6). Though we found no variation in the copy number of any of these genes among the three sequenced reference genomes included in the *Eubacterium rectale* cluster in our database, a recent study of 27 elderly gut metagenomes identified non-uniform coverage of genes involved in the flagellum biogenesis pathways of six *Eubacterium* and *Roseburia* species (Neville et al., 2013), suggesting that the current reference genomes may not capture the full dynamic range of these species.

Next, we considered the collection of set-specific variable KCs and examined their functional annotations. Interestingly, hierarchical clustering of set-specific variable KOs based solely on their variation profile across the 40 clusters revealed distinct groups of functionally related genes that vary in a given genome cluster or within multiple clusters (Figure 3). For example, a large set of genes related to cell growth and



**Figure 5. Copy Number of Highly Variable Transport KCs in *Bacteroides ovatus***

The size and color of each square represent the copy number of each highly variable KC within each sample. Samples are grouped by host state (I, IBD; h, healthy; o, obese). The copy numbers of the 13 marker KCs in this genome cluster are illustrated for comparison. See also [Figure S6](#).

sporulation were all identified as set-specific variable KCs in the two genome clusters associated with *Clostridium* sp. Similarly, a set of sugar metabolism genes were all identified as set-specific variable KCs in *Roseburia intestinalis*, and a number of antibiotic resistance genes were identified as variable in multiple genome clusters, primarily those in the *Firmicutes* phylum. An enrichment analysis of functions associated with set-specific variable KCs in each cluster additionally revealed a number of important functions that were prone to copy-number variation ([Table S6](#)). For example, genes in the lipopolysaccharide biosynthesis pathway in *Dialister invisus* and *Clostridium* sp. were often observed with a higher copy number in a small set of samples. Interestingly, variation within functions related to sugar metabolism (i.e., KEGG pathways galactose metabolism, starch and sucrose metabolism, fructose and mannose metabolism, polyketide sugar unit biosynthesis) was observed primarily within *Bacteroidetes* clusters, whereas set-specific transport-related variation was almost absent from these clusters. Other functions enriched for set-specific variable KCs suggest transitions between virulent states, such as motility in *butyrate-producing bacteria* (NCBI accession FP929062), *Eubacterium rectale*, and *Clostridium* sp.; streptomycin biosynthesis in *Acidaminococcus* sp.; lysosome production in *Bacteroides ovatus*; the EHEC/EPEC pathogenicity signature in *Escherichia coli*; and secretion systems in *butyrate-producing bacteria* (NCBI accession FP929062), *Clostridium* sp., and *Escherichia coli*. Within *Escherichia coli*, type II secretion system genes were identified as set-specific decreased copy-number KCs, whereas type III secretion system genes were identified as set-specific increased copy-number KCs. Overall, much of the observed variation appeared to be associated with the way a species responds to and interacts with its surroundings, highlighting the strong adaptive potential of gut-associated bacteria.

Clearly, different cohorts could harbor different sets of strains owing to an assortment of ecological or host-specific factors, and accordingly different genes may vary in copy number in

different datasets. Notably, however, analysis of a second dataset of 73 gut samples from a Chinese cohort ([Qin et al., 2012](#)) yielded a marked overlap with our original Danish/Spanish cohort in both the set of KCs identified as variable and the set of functions enriched for copy-number variation ([Extended Experimental Procedures](#)). These findings suggest that, although variation may be personal, certain genes and functions (e.g., those related to environmental adaptation) may be universally prone to variation.

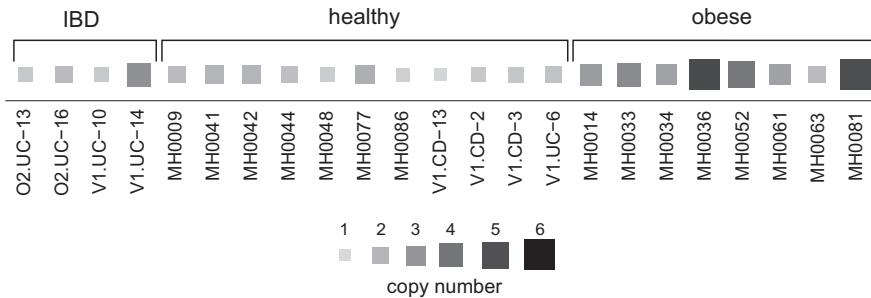
### Host State-Associated Variation

Although much of the variation across strains may reflect neutral processes or transitory dynamics, some variation may represent adaptation to a specific host phenotype. To detect such potentially adaptive variation, we identified variable KCs in which the copy number in samples from obese or IBD subjects was significantly different than in samples from healthy subjects ([Experimental Procedures](#)). In total, we found 24 KCs whose copy number was significantly associated with IBD and three KCs whose copy number was significantly associated with obesity (FDR < 0.05; [Table S7](#)).

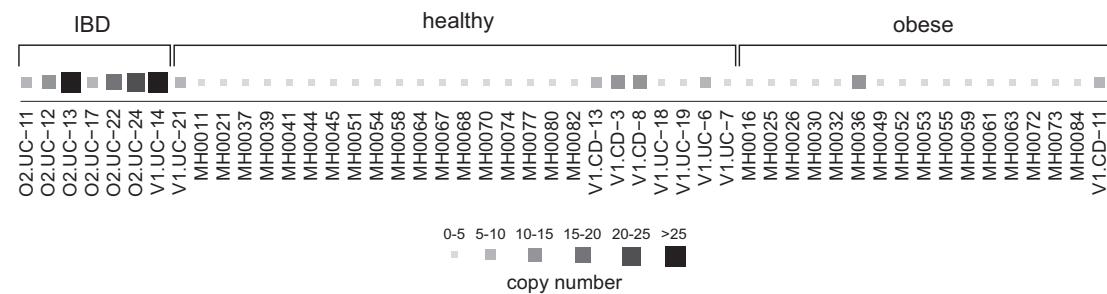
Interestingly, a number of these KCs have been previously implicated in adverse host health states. For example, in our analysis, obesity was associated with a higher copy number of thioredoxin 1 (K03671) in *Clostridium* sp. ([Figure 6A](#)), and indeed thioredoxin reductase was recently shown to be enriched in the cecal metaproteome of mice fed a high-fat diet ([Daniel et al., 2014](#)). Such results are consistent with thioredoxin's regulatory role in maintaining redox equilibrium and the demonstrated links between a high-fat diet and oxidative stress in mammals ([Furukawa et al., 2004](#)). Additionally, in our analysis, the loss of a ubiquinone-reducing gene (K00349; *nqrD*) from *Bacteroides plebeius* was associated with obesity. A recent study in mice showed that supplemental ubiquinone reduced inflammation and metabolic stress accompanying a high-fat high-fructose diet by reducing the expression of certain genes associated

**A**

K03671 (thioredoxin 1) in cluster 49 (Clostridium sp.)

**B**

K08217 (MFS transporter, macrolide efflux) in cluster 20 (Roseburia inulinivorans)

**Figure 6. Copy-Number Variation of Host State-Associated KCs**

Two KCs whose copy number was significantly increased in samples from a specific host state are shown. The size and color of each square represent the copy number of the KC within each sample.

(A) The copy number of thioredoxin 1 (K03671) in Clostridium sp. is significantly increased in samples from obese subjects.

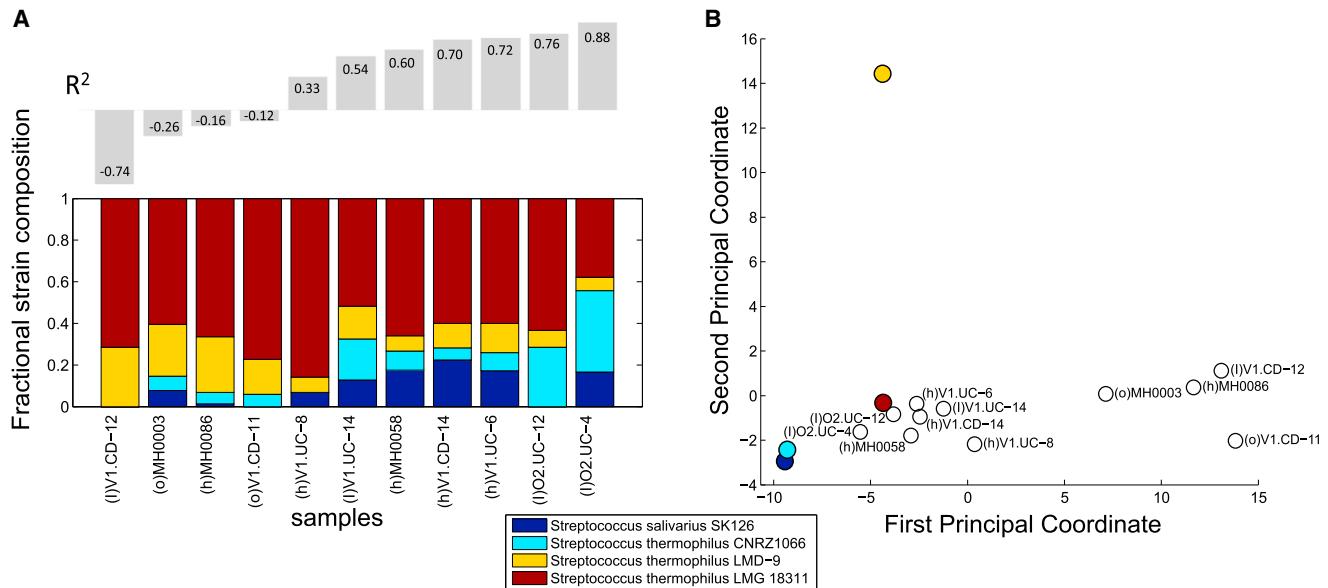
(B) The copy number of an MFS transporter gene (K08217) in the Roseburia inulinivorans genome cluster is significantly increased in samples from IBD subjects. See also Table S7.

with stress-response (Sohet et al., 2009), while mice not receiving the supplement gained more weight than their counterparts. Importantly, however, ubiquinol, the reduced form of ubiquinone, has recently been shown to be the more readily absorbed and more active form of the compound (Langsjoen and Langsjoen, 2014), raising the possibility that loss of microbial ubiquinone-reducing capabilities from certain species may hinder the effectiveness and protective capacity of ubiquinone in the host. Other findings shed new light on the role of individual species in disease, with evidence of variation associated with common disease hallmarks, such as pathogenicity-related secretion and antibiotic resistance. In *Roseburia inulinivorans* (Figure 6B), increased copy number of a gene (K08217) coding for a major drug efflux protein known to play a role in antibiotic resistance was highly associated with IBD-afflicted individuals. Similarly, HlyD (K01993), an essential component of RTX hemolytic toxin secretion (Pimenta et al., 2005), exhibited increased copy number in IBD samples in *Bacteroides uniformis*. See Table S7 for a full list of disease-associated KCs. Interestingly, none of the obesity-associated KCs and only 3 of the 24 IBD-associated KCs were found to vary significantly in the Chinese cohort described above, among whom only one individual was obese and none were reported as having IBD.

### Strain-Level Deconvolution of Microbiome Composition and Intra-Species Population Structure

Clearly, the microbiomes of different individuals can house multiple strains of the same species with potentially different relative abundances. Our copy-number estimates for each cluster accordingly represent average copy numbers across the different strains in the sample. Next, we therefore examined whether these estimates can be used to obtain insights into strain-level population structure, going beyond species-level composition assays and focusing specifically on the composition of strains within each genome cluster rather than on the abundance of the cluster itself.

First, we explored how well the copy-number profiles obtained for each genome cluster in each sample can be explained by known reference strains, using a regression analysis to deconvolve these copy-number profiles into a linear combination of the strains included in our database (Experimental Procedures). Obviously, these strains may not encompass the full set of strains present in the samples analyzed, yet such an analysis may be useful in examining what portion of the observed variation can be accounted for by known strains and what portion represents potentially novel variation. Indeed, we found that, in well-characterized clusters with many sequenced genomes,



**Figure 7. Predicted Strain-Level Population Structure within *Clostridium* sp**

(A) A linear regression analysis was used to model the copy-number profile obtained for cluster 110 (*Streptococcus thermophilus*) in each sample as a combination of known reference genomes, with prediction weights shown as stacked colored bars. Prediction accuracy ( $R^2$ ) is indicated above each bar. Samples with low or negative  $R^2$  values potentially contain variation that cannot be explained by any combination of known reference genomes.

(B) A principal coordinate analysis depicting the differences between the copy-number profiles obtained for this genome cluster in the various samples (open circles), as well as the copy-number profiles of reference genomes (filled circles). See also Figure S7.

the copy-number profiles of most samples could be well explained by a linear combination of known strains. For example, in the *Escherichia coli* cluster that comprised 63 sequenced genomes in our database, 76% of the variation in copy number could be explained on average by these genomes ( $R^2 = 0.76 \pm 0.12$ ). In this cluster, the inferred representation of each strain differed widely across samples, with some strains (i.e., *Escherichia coli* O111:H- str. 11128) highly represented across multiple samples and others found in just one sample. However, for less well-characterized clusters with only a few known strains in our database, in some cases just a subset of the observed copy-number variation could be explained. For example, the four known strains of *Streptococcus thermophilus* could be used to explain a majority of the variation observed in some of the samples ( $R^2 > 0.5$ ) yet failed to explain the variation observed in four of the samples ( $R^2 < 0$ ), suggesting the existence of potentially novel, yet-to-be-sequenced variation (Figure 7A).

To further compare copy-number variation profiles across samples and to examine variation that may not be captured by known strains (including notably, in clusters comprising only one known strain), we used a principal coordinate analysis. This analysis revealed a complex population structure within each cluster, with marked differences among samples indicating the prevalence of personalized variation. For a number of genome clusters, however, samples appear to group into distinct sets, potentially reflecting individuals with similar intra-species population structures (Figure 7B). Moreover, by including the reference genomes in this principal coordinate analysis, we were able to distinguish previously captured

variation versus novel variation observed across samples. For example, the principal coordinate plot for the *Streptococcus thermophilus* genome cluster (Figure 7B) clearly demonstrates that, although the copy-number profiles of most samples clustered tightly with several known reference genomes, the four poorly explained samples mentioned above clustered together and contained variation that was distinct from any reference genome. Such a pattern may indicate the presence of novel shared strains, providing a promising basis for targeted sequencing. Similar patterns were also observed in other clusters, in which a distinct, tightly clustered subset of samples or individual samples exhibit markedly different copy-number profile from that of any sequenced genome (Figures S7A and S7B). Overall though, each genome cluster exhibited a unique population structure across individuals, highlighting the complex suite of forces governing taxonomic composition in the gut (Levy and Borenstein, 2013).

## DISCUSSION

By and large, closely related organisms tend to encode similar sets of genes. This consistency is in fact often used to infer functional capacity from taxonomy (Langille et al., 2013; Zaneveld et al., 2010). Clearly, however, this relationship between phylogeny and gene content is imperfect, and each species represents a large collection of strains that differ in the set of genes they encode, the copy number of these genes, and ultimately, their functional capacity. Above, we have focused on identifying instances in which this relationship between microbial species

and genes breaks, presenting a large-scale analysis of copy-number variation in a diverse array of gut species. Our analysis has demonstrated that copy-number variation is prevalent in the gut environment, with some species exhibiting significant copy-number variation in >20% of their genes. Such variation may induce significant microbiome-wide shifts and may account for at least some of the observed discrepancies between trends observed at the species levels versus trends measured at the gene level. Moreover, intra-species variation was shown to be especially prevalent in genes involved in specific functions, most notably functions that impact the way an organism interacts with its environment such as transport and signaling processes. This may suggest an adaptive dynamic by which certain species respond to changes in community composition or in the gut niche and a potentially crucial role of the gut environment in shaping bacterial evolution (Levy and Borenstein, 2013; Shapiro et al., 2012). Other highly variable functions, such as lipopolysaccharide biosynthesis, cell motility, and secretion systems, may represent changes in virulence as organisms respond to host immune responses. Interestingly, many of these same functions were highlighted in a previous study as more difficult to accurately correlate with 16s data (Langille et al., 2013). Our analysis further identified variable functions that may correlate with host states, exhibiting differential copy number in specific genomes. It remains unclear, however, whether such host state-associated variation is a cause or an effect. Our framework additionally facilitated the inference of intra-species population profiles for each individual, suggesting that most individuals harbor multiple strains of each species.

Though still far from an exhaustive catalog of strains that may be present across all human gut microbiomes, the framework presented above represents the most comprehensive account of copy-number variation in the human gut microbiome to date. It is our hope that this framework and the results presented here will inform future studies of strain-level microbiome composition, demonstrating the extent of functional information that is lost by limiting characterization to the level of species and prompting further investigation and sequencing of strain-level features. Yet, there are clearly a number of caveats that should be considered in designing such future efforts. First, our analysis is limited to the detection of variation in gut species for which at least one fully sequenced genome is available, and future studies may benefit from additional genomes. Notably though, we did not detect significantly more variation in clusters for which more reference genomes were available. In addition, our pipeline was designed to detect gene losses or amplifications but cannot identify gain of genes that are not present in any of the reference genomes included in the genome cluster. Such gain or transfer events may represent an additional substantial source of intra-species variation (Smillie et al., 2011). Our framework could, however, further facilitate future efforts to study sequence divergence among duplicated genes, informing our view of neo-functionalization and conservation processes in the microbiome. Notably, in our analysis, we focused on detecting high-confidence instances of variation, applying conservative parameters for read alignment and for variability calling. Specifically, we limit our analysis to “detectable” genome clusters, defined as those with >1× coverage in the sample. Our analysis of a synthetic da-

taset confirmed that, in such clusters, copy-number estimates can be inferred with 96% accuracy but that prediction accuracy dropped significantly in genome clusters with lower coverage (Figure S4B and Extended Experimental Procedures). With 13 million reads per sample (the lowest sequencing depth in the cohort analyzed), species that comprise >0.4% of the sample are likely to be considered detectable by our pipeline (while a higher sequencing depth of a sample will clearly allow analysis of even rarer species). Future studies may relax some of these parameters or incorporate additional information (e.g., gene conservation) to detect more subtle variation. Finally, as with most studies relating microbiome composition to function, our analysis relies on the availability of functional gene databases, which may contain incomplete or erroneous annotations. By considering variation across samples rather than variation from reference genomes, our analysis is largely robust to such annotation inaccuracies. Interestingly, however, variable KCs identified by our analysis were much more likely to lack a functional annotation than non-variable KCs, suggesting that much of the detected variation in gene content has as yet uncharacterized consequences. Combined, these results highlight both the need for additional genome sequences and the importance of continued efforts for characterizing gene function.

Ultimately, analysis of intra-species variation in microbial communities is crucial for understanding the complex relationship between species composition and community-level functional capacity. Our analysis, quantifiably characterizing such variation in the gut microbiome, is an important first step in this direction, and the resulting dataset provides an essential resource for future predictive studies.

## EXPERIMENTAL PROCEDURES

### Metagenomic Samples and Reference Genomes

Gut metagenomic data for 109 Danish and Spanish individuals, including individuals afflicted with obesity or IBD, was obtained from (Qin et al., 2010). A list of 261 dominant and prevalent human gut microbial strains, grouped into 101 genome clusters (Table S1) based on sequence similarity of 40 marker genes, was obtained from (Schloissnig et al., 2013). Nucleotide contig sequences, gene calls, and amino acid protein sequences were downloaded for each genome, and protein sequences were annotated with KEGG orthologous groups (KOs). See Extended Experimental Procedures for more details.

### Calculation of Copy-Number Estimates

Shotgun metagenomic reads were aligned to the set of reference genomes with BWA, using parameters and filters carefully validated by extensive simulation analyses (Figures S1 and S2; Extended Experimental Procedures). In total, 2,469,102,286 reads were mapped. Average coverage over each gene region was determined using samtools (Li et al., 2009), and the coverage of each KC (KO-cluster pair) was obtained by summing over all genes annotated with the same KO and genome cluster. KC coverage was normalized by cluster abundance, defined as the average coverage over a set of 13 universal marker KOs (Figure S3B; Extended Experimental Procedures), to obtain the estimated copy number  $V_{kcs}$  of each KO  $k$ , in each cluster  $c$ , and in each sample  $s$ . “Detectable KCs” in a sample were defined as those with  $V_{kcs} \geq 0.5$ . “Detectable clusters” within each sample were defined as those with at least 12 detectable marker KCs and average marker coverage  $\geq 1$ . KCs that were not detectable in any sample were removed from the analysis.

### Detection of Highly Variable and Set-Specific Variable KCs

For each of the 40,088 KCs present in clusters detectable in at least ten samples, the median copy number (baseline) across samples and the MAD

(median absolute deviation) from this baseline were calculated. KCs with a MAD more than 2 SDs from the MAD distribution mean ( $MAD > 0.6346$ ) were considered *highly variable*. KCs in which at least 10% of samples had a copy number that exceeded the baseline by this threshold were considered *set-specific increased variable KCs*. *Set-specific decreased KCs* were similarly defined as KCs in which at least 10% of samples had a copy number that fell below the baseline by this threshold.

#### Detection of Host State-Associated KCs

A KC was defined as obesity associated if the copy numbers in samples from obese individuals were significantly higher or significantly lower than the copy numbers in samples from non-obese individuals, according to a two-sample t test (FDR-corrected  $p < 0.05$ ). IBD-associated KCs were similarly defined. Samples that were labeled as both obese and IBD were omitted from this analysis.

#### Copy-Number Profile Deconvolution and Principal Coordinate Analysis

For each sample, a non-negative least-squares linear regression analysis was performed to obtain the linear combination of reference genomes in each multi-genome cluster, optimally explaining the copy-number estimates of variable KCs. The regression was constrained such that the sum of genome weights for each sample and cluster equaled one. Prediction error was defined as the  $R^2$  value for each sample. A principal coordinate analysis was also performed for every genome cluster, operating on the pairwise Euclidian distance matrix of set-specific variable KC copy numbers in each sample and each sequenced reference genome.

#### SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, seven figures, and seven tables and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2014.12.038>.

#### ACKNOWLEDGMENTS

We thank Peter Turnbaugh, Jay Shendure, Phil Green, Colin Manoil, two anonymous reviewers, and the members of the Borenstein Lab for support and helpful discussions. This work was supported by a New Innovator Award DP2 AT 007802-01 to EB.

Received: June 19, 2014

Revised: October 30, 2014

Accepted: December 24, 2014

Published: January 29, 2015

#### REFERENCES

Al Mamun, A.A., Tominaga, A., and Enomoto, M. (1997). Cloning and characterization of the region III flagellar operons of the four *Shigella* subgroups: genetic defects that cause loss of flagella of *Shigella boydii* and *Shigella sonnei*. *J. Bacteriol.* **179**, 4493–4500.

Borziak, K., Fleetwood, A.D., and Zhulin, I.B. (2013). Chemoreceptor gene loss and acquisition via horizontal gene transfer in *Escherichia coli*. *J. Bacteriol.* **195**, 3596–3602.

Brown, C.T., Sharon, I., Thomas, B.C., Castelle, C.J., Morowitz, M.J., and Banfield, J.F. (2013). Genome resolved analysis of a premature infant gut microbial community reveals a *Varibaculum cambriense* genome and a shift towards fermentation-based metabolism during the third week of life. *Microbiome* **1**, 30.

Daniel, H., Moghaddas Gholami, A., Berry, D., Desmarchelier, C., Hahne, H., Loh, G., Mondot, S., Lepage, P., Rothbauer, M., Walker, A., et al. (2014). High-fat diet alters gut microbiota physiology in mice. *ISME J.* **8**, 295–308.

Fitzsimons, M.S., Novotny, M., Lo, C.-C., Dichosa, A.E.K., Yee-Greenbaum, J.L., Snook, J.P., Gu, W., Chertkov, O., Davenport, K.W., McMurry, K., et al. (2013). Nearly finished genomes produced using gel microdroplet culturing reveal substantial intraspecies genomic diversity within the human microbiome. *Genome Res.* **23**, 878–888.

Frank, D.N., St Amand, A.L., Feldman, R.A., Boedeker, E.C., Harpaz, N., and Pace, N.R. (2007). Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. *Proc. Natl. Acad. Sci. USA* **104**, 13780–13785.

Furukawa, S., Fujita, T., Shimabukuro, M., Iwaki, M., Yamada, Y., Nakajima, Y., Nakayama, O., Makishima, M., Matsuda, M., and Shimomura, I. (2004). Increased oxidative stress in obesity and its impact on metabolic syndrome. *J. Clin. Invest.* **114**, 1752–1761.

Gevers, D., Vandepoele, K., Simillon, C., and Van de Peer, Y. (2004). Gene duplication and biased functional retention of paralogs in bacterial genomes. *Trends Microbiol.* **12**, 148–154.

Gill, S.R., Fouts, D.E., Archer, G.L., Mongodin, E.F., Debroy, R.T., Ravel, J., Paulsen, I.T., Kolonay, J.F., Brinkac, L., Beanan, M., et al. (2005). Insights on evolution of virulence and resistance from the complete genome analysis of an early methicillin-resistant *Staphylococcus aureus* strain and a biofilm-producing methicillin-resistant *Staphylococcus epidermidis* strain. *J. Bacteriol.* **187**, 2426–2438.

Han, X., Kennan, R.M., Davies, J.K., Reddacliff, L.A., Dhungel, O.P., Whittington, R.J., Turnbull, L., Whitchurch, C.B., and Rood, J.I. (2008). Twitching motility is essential for virulence in *Dichelobacter nodosus*. *J. Bacteriol.* **190**, 3323–3335.

Hansen, E.E., Lozupone, C.A., Rey, F.E., Wu, M., Guruge, J.L., Narra, A., Goodfellow, J., Zaneveld, J.R., McDonald, D.T., Goodrich, J.A., et al. (2011). Pan-genome of the dominant human gut-associated archaeon, *Methanobrevibacter smithii*, studied in twins. *Proc. Natl. Acad. Sci. USA* **108** (1), 4599–4606.

Heikkinen, E., Kallonen, T., Saarinen, L., Sara, R., King, A.J., Mooi, F.R., Soini, J.T., Mertsola, J., and He, Q. (2007). Comparative genomics of *Bordetella pertussis* reveals progressive gene loss in Finnish strains. *PLoS ONE* **2**, e904.

Hoffman, L.R., Pope, C.E., Hayden, H.S., Heltshe, S., Levy, R., McNamara, S., Jacobs, M.A., Rohmer, L., Radey, M., Ramsey, B.W., et al. (2014). *Escherichia coli* dysbiosis correlates with gastrointestinal dysfunction in children with cystic fibrosis. *Clin. Infect. Dis.* **58**, 396–399.

Human Microbiome Project Consortium (2012). Structure, function and diversity of the healthy human microbiome. *Nature* **486**, 207–214.

Iida, N., Dzutsev, A., Stewart, C.A., Smith, L., Bouladoux, N., Weingarten, R.A., Molina, D.A., Salcedo, R., Back, T., Cramer, S., et al. (2013). Commensal bacteria control cancer response to therapy by modulating the tumor microenvironment. *Science* **342**, 967–970.

Kinross, J.M., Darzi, A.W., and Nicholson, J.K. (2011). Gut microbiome-host interactions in health and disease. *Genome Med.* **3**, 14.

Kraal, L., Abubucker, S., Kota, K., Fischbach, M.A., and Mitreva, M. (2014). The prevalence of species and strains in the human microbiome: a resource for experimental efforts. *PLoS ONE* **9**, e97279.

Langille, M.G.I., Zaneveld, J., Caporaso, J.G., McDonald, D., Knights, D., Reyes, J.A., Clemente, J.C., Burkepile, D.E., Vega Thurber, R.L., Knight, R., et al. (2013). Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat. Biotechnol.* **31**, 814–821.

Langsjoen, P.H., and Langsjoen, A.M. (2014). Comparison study of plasma coenzyme Q 10 levels in healthy subjects supplemented with ubiquinol versus ubiquinone. *Clin. Pharmacol. Drug Dev.* **3**, 13–17.

Larsen, N., Vogensen, F.K., van den Berg, F.W.J., Nielsen, D.S., Andreasen, A.S., Pedersen, B.K., Al-Soud, W.A., Sørensen, S.J., Hansen, L.H., and Jakobsen, M. (2010). Gut microbiota in human adults with type 2 diabetes differs from non-diabetic adults. *PLoS ONE* **5**, e9085.

Lee, M.-C., and Marx, C.J. (2012). Repeated, selection-driven genome reduction of accessory genes in experimental populations. *PLoS Genet.* **8**, e1002651.

Levy, R., and Borenstein, E. (2013). Metabolic modeling of species interaction in the human microbiome elucidates community-level assembly rules. *Proc. Natl. Acad. Sci. USA* **110**, 12804–12809.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079.

Morowitz, M.J., Denef, V.J., Costello, E.K., Thomas, B.C., Poroyko, V., Relman, D.A., and Banfield, J.F. (2011). Strain-resolved community genomic analysis of gut microbial colonization in a premature infant. *Proc. Natl. Acad. Sci. USA* 108, 1128–1133.

Muegge, B.D., Kuczynski, J., Knights, D., Clemente, J.C., González, A., Fontana, L., Henrissat, B., Knight, R., and Gordon, J.I. (2011). Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. *Science* 332, 970–974.

Neville, B.A., Sheridan, P.O., Harris, H.M.B., Coughlan, S., Flint, H.J., Duncan, S.H., Jeffery, I.B., Claesson, M.J., Ross, R.P., Scott, K.P., and O'Toole, P.W. (2013). Pro-inflammatory flagellin proteins of prevalent motile commensal bacteria are variably abundant in the intestinal microbiome of elderly humans. *PLoS ONE* 8, e68919.

Pimenta, A.L., Racher, K., Jamieson, L., Blight, M.A., and Holland, I.B. (2005). Mutations in HlyD, part of the type 1 translocator for hemolysin secretion, affect the folding of the secreted toxin. *J. Bacteriol.* 187, 7471–7480.

Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K.S., Manichanh, C., Nielsen, T., Pons, N., Levenez, F., Yamada, T., et al.; MetaHIT Consortium (2010). A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* 464, 59–65.

Qin, J., Li, Y., Cai, Z., Li, S., Zhu, J., Zhang, F., Liang, S., Zhang, W., Guan, Y., Shen, D., et al. (2012). A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature* 490, 55–60.

Salama, N., Guillemin, K., McDaniel, T.K., Sherlock, G., Tompkins, L., and Falkow, S. (2000). A whole-genome microarray reveals genetic diversity among *Helicobacter pylori* strains. *Proc. Natl. Acad. Sci. USA* 97, 14668–14673.

Schloissnig, S., Arumugam, M., Sunagawa, S., Mitreva, M., Tap, J., Zhu, A., Waller, A., Mende, D.R., Kultima, J.R., Martin, J., et al. (2013). Genomic variation landscape of the human gut microbiome. *Nature* 493, 45–50.

Shapiro, B.J., Friedman, J., Cordero, O.X., Preheim, S.P., Timberlake, S.C., Szabó, G., Polz, M.F., and Alm, E.J. (2012). Population genomics of early events in the ecological differentiation of bacteria. *Science* 336, 48–51.

Sharon, I., Morowitz, M.J., Thomas, B.C., Costello, E.K., Relman, D.A., and Banfield, J.F. (2013). Time series community genomics analysis reveals rapid shifts in bacterial species, strains, and phage during infant gut colonization. *Genome Res.* 23, 111–120.

Siezen, R.J., Tzeneva, V.A., Castioni, A., Wels, M., Phan, H.T.K., Rademaker, J.L.W., Starrenburg, M.J.C., Kleerebezem, M., Molenaar, D., and van Hyckama Vlieg, J.E.T. (2010). Phenotypic and genomic diversity of *Lactobacillus plantarum* strains isolated from various environmental niches. *Environ. Microbiol.* 12, 758–773.

Smillie, C.S., Smith, M.B., Friedman, J., Cordero, O.X., David, L.A., and Alm, E.J. (2011). Ecology drives a global network of gene exchange connecting the human microbiome. *Nature* 480, 241–244.

Sohet, F.M., Neyrinck, A.M., Pachikian, B.D., de Backer, F.C., Bindels, L.B., Niklowitz, P., Menke, T., Cani, P.D., and Delzenne, N.M. (2009). Coenzyme Q10 supplementation lowers hepatic oxidative stress and inflammation associated with diet-induced obesity in mice. *Biochem. Pharmacol.* 78, 1391–1400.

Solheim, M., Aakra, A., Snipen, L.G., Brede, D.A., and Nes, I.F. (2009). Comparative genomics of *Enterococcus faecalis* from healthy Norwegian infants. *BMC Genomics* 10, 194.

Sonti, R.V., and Roth, J.R. (1989). Role of gene duplications in the adaptation of *Salmonella typhimurium* to growth on limiting carbon sources. *Genetics* 123, 19–28.

Turnbaugh, P.J., Hamady, M., Yatsunenko, T., Cantarel, B.L., Duncan, A., Ley, R.E., Sogin, M.L., Jones, W.J., Roe, B.A., Affourtit, J.P., et al. (2009). A core gut microbiome in obese and lean twins. *Nature* 457, 480–484.

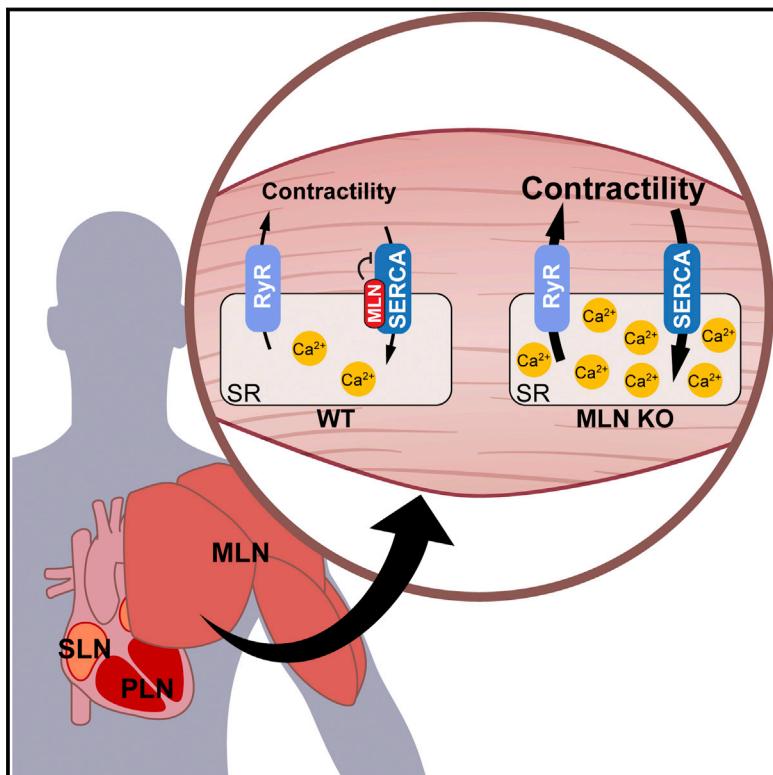
Vijay-Kumar, M., Aitken, J.D., Carvalho, F.A., Cullender, T.C., Mwangi, S., Srivivasan, S., Sitaraman, S.V., Knight, R., Ley, R.E., and Gewirtz, A.T. (2010). Metabolic syndrome and altered gut microbiota in mice lacking Toll-like receptor 5. *Science* 328, 228–231.

Zaneveld, J.R., Lozupone, C., Gordon, J.I., and Knight, R. (2010). Ribosomal RNA diversity predicts genome diversity in gut bacteria and their relatives. *Nucleic Acids Res.* 38, 3869–3879.

Zunino, P., Piccini, C., and Legnani-Fajardo, C. (1994). Flagellate and non-flagellate *Proteus mirabilis* in the development of experimental urinary tract infection. *Microb. Pathog.* 16, 379–385.

# A Micropeptide Encoded by a Putative Long Noncoding RNA Regulates Muscle Performance

## Graphical Abstract



## Authors

Douglas M. Anderson,  
Kelly M. Anderson, ...,  
Rhonda Bassel-Duby, Eric N. Olson

## Correspondence

eric.olson@utsouthwestern.edu

## In Brief

Myoregulin is a skeletal muscle-specific micropeptide that regulates muscle performance by modulating intracellular calcium handling.

## Highlights

- Myoregulin is a micropeptide encoded by an annotated long noncoding RNA
- Myoregulin is a transmembrane alpha helix expressed only in skeletal muscle
- Myoregulin regulates Ca<sup>2+</sup> handling by inhibiting the pump activity of SERCA
- Myoregulin KO mice show improved exercise performance and Ca<sup>2+</sup> handling in muscle

# A Micropeptide Encoded by a Putative Long Noncoding RNA Regulates Muscle Performance

Douglas M. Anderson,<sup>1,4</sup> Kelly M. Anderson,<sup>1,4</sup> Chi-Lun Chang,<sup>2</sup> Catherine A. Makarewicz,<sup>1,4</sup> Benjamin R. Nelson,<sup>1,4</sup> John R. McAnally,<sup>1,4</sup> Prasad Kasaragod,<sup>1</sup> John M. Shelton,<sup>3</sup> Jen Liou,<sup>2</sup> Rhonda Bassel-Duby,<sup>1,4</sup> and Eric N. Olson<sup>1,4,\*</sup>

<sup>1</sup>Department of Molecular Biology

<sup>2</sup>Department of Physiology

<sup>3</sup>Department of Internal Medicine

<sup>4</sup>Hamon Center for Regenerative Science and Medicine

The University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX 75390–9148, USA

\*Correspondence: eric.olson@utsouthwestern.edu

<http://dx.doi.org/10.1016/j.cell.2015.01.009>

## SUMMARY

Functional micropeptides can be concealed within RNAs that appear to be noncoding. We discovered a conserved micropeptide, which we named myoregulin (MLN), encoded by a skeletal muscle-specific RNA annotated as a putative long noncoding RNA. MLN shares structural and functional similarity with phospholamban (PLN) and sarcolipin (SLN), which inhibit SERCA, the membrane pump that controls muscle relaxation by regulating  $\text{Ca}^{2+}$  uptake into the sarcoplasmic reticulum (SR). MLN interacts directly with SERCA and impedes  $\text{Ca}^{2+}$  uptake into the SR. In contrast to PLN and SLN, which are expressed in cardiac and slow skeletal muscle in mice, MLN is robustly expressed in all skeletal muscle. Genetic deletion of MLN in mice enhances  $\text{Ca}^{2+}$  handling in skeletal muscle and improves exercise performance. These findings identify MLN as an important regulator of skeletal muscle physiology and highlight the possibility that additional micropeptides are encoded in the many RNAs currently annotated as noncoding.

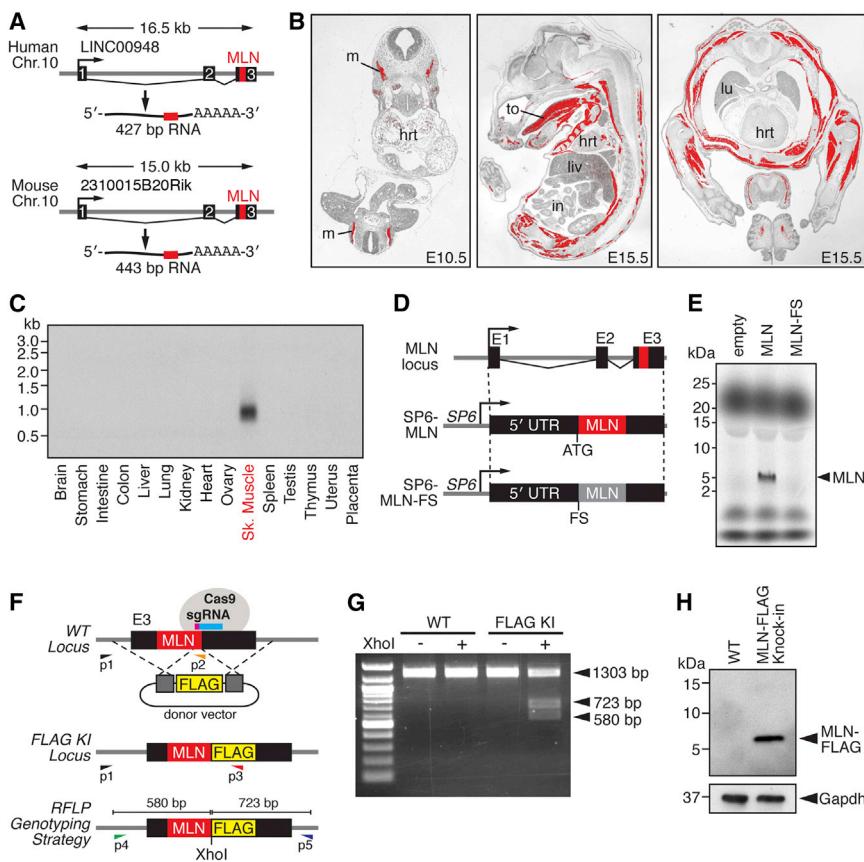
## INTRODUCTION

$\text{Ca}^{2+}$  controls the normal function of striated muscle by acting as the primary regulator of the sarcomeric contractile machinery and as a second messenger in the signal transduction pathways that control muscle growth, metabolism, and pathological remodeling (Bassel-Duby and Olson, 2006; Berchtold et al., 2000).  $\text{Ca}^{2+}$  handling in striated muscle is tightly regulated by  $\text{Ca}^{2+}$  pumps in the sarcoplasmic reticulum (SR) and plasma membranes that maintain intracellular  $\text{Ca}^{2+}$  levels ~10,000-fold lower than extracellular and SR concentrations (Berridge et al., 2003; Rossi and Dirksen, 2006). Upon muscle stimulation,  $\text{Ca}^{2+}$  release by the ryanodine receptor (RyR) in the SR membrane transiently increases  $\text{Ca}^{2+}$  levels in the cytosol, triggering actomyosin cross-bridge formation within the sarcomere to generate contractile force. Reuptake of  $\text{Ca}^{2+}$  into the SR by sarcoplasmic reticulum

$\text{Ca}^{2+}$ -ATPase (SERCA) is necessary for muscle relaxation and restores SR  $\text{Ca}^{2+}$  levels for subsequent contraction-relaxation cycles. SERCA serves as a central regulator of striated muscle performance and the pathological signaling pathways that drive cardiovascular and skeletal muscle disease (Dorn and Molkentin, 2004; Goonasekera et al., 2011; Odermatt et al., 1996; Pan et al., 2003; Periasamy and Kalyanasundaram, 2007).

Two related peptides, phospholamban (PLN) and sarcolipin (SLN), directly interact with SERCA in the SR membrane to regulate  $\text{Ca}^{2+}$  pump activity (Kranias and Hajjar, 2012; MacLennan and Kranias, 2003). PLN and SLN are expressed in partially overlapping patterns in cardiac and slow skeletal muscle and are important regulators of muscle performance and cardiovascular disease (Briggs et al., 1992; Kranias and Hajjar, 2012; Minamisawa et al., 2003; Schmitt et al., 2003; Tada and Toyofuku, 1998; Tupling et al., 2011). PLN-deficient mice exhibit enhanced myocardial contractile performance, characterized by increased ventricular relaxation rates and SERCA pump activity (Chu et al., 1998; Luo et al., 1994). Similarly, loss of PLN or SLN expression significantly increases the rate of muscle relaxation and SERCA pump activity in slow skeletal muscle but does not affect fast skeletal muscles, which do not express PLN or SLN (Slack et al., 1997; Tupling et al., 2011; Vangheluwe et al., 2005). The absence of PLN and SLN expression in fast skeletal muscle, the dominant muscle type in mice, suggests that an unidentified factor regulates  $\text{Ca}^{2+}$  handling and the contractile performance of this tissue.

Recent genome-wide studies have suggested that hundreds of functional micropeptides may be encoded in vertebrate long noncoding RNAs (lncRNAs) (Andrews and Rothnagel, 2014; Bazzini et al., 2014). The microproteome has largely been overlooked in gene annotations, primarily because of the difficulty in identifying functional small open reading frames (ORFs) in RNA transcripts. While analyzing an annotated skeletal muscle-specific lncRNA, we discovered a previously unrecognized ORF encoding a conserved 46 amino acid micropeptide, which we named myoregulin (MLN). MLN forms a single transmembrane alpha helix that interacts with SERCA in the membrane of the SR and regulates  $\text{Ca}^{2+}$  handling. Consistent with this function, deletion of MLN in mice significantly enhances  $\text{Ca}^{2+}$  handling and improves exercise performance. These findings identify MLN as the predominant SERCA-inhibitory



(G) RFLP analysis of WT C2C12 and heterozygous C2C12 myoblasts for the MLN-FLAG knockin allele.

(H) Western blot analysis showing endogenous expression of the MLN-FLAG fusion peptide in differentiated C2C12 myotubes, detected with an anti-FLAG antibody.

See also Figure S1.

micropeptide in skeletal muscle, which surprisingly was concealed in an RNA annotated as noncoding.

## RESULTS

### Discovery of a Conserved Micropeptide Encoded by a lncRNA

In a bioinformatic screen for uncharacterized skeletal muscle-specific genes, we identified a vertebrate RNA transcript annotated as a lncRNA (LINC00948 in humans and 2310015B20Rik in mice). Analysis of the evolutionary conservation of these transcripts identified a short 138 nucleotide ORF with the potential to encode a highly conserved 46 amino acid micropeptide, which we named myoregulin (MLN) (Figure S1A available online). The human and mouse MLN genes consist of three exons that span 16.5 and 15.0 kb, respectively, with the ORF located in exon 3 (Figure 1A). Nucleotide insertions and deletions that could alter the reading frame flank the MLN ORF, demonstrating that these sequences comprise UTRs (Figure S1A).

During embryogenesis, MLN is expressed in the myotomal compartment of the somites, the anlagen of skeletal muscle (Figure 1B). During fetal and adult stages, MLN is robustly expressed in all skeletal muscles and is not detectable in cardiac or smooth

### Figure 1. Discovery of a Skeletal Muscle-Specific Micropeptide

(A) A short ORF encoding a conserved micropeptide, which we named myoregulin (MLN), is contained within exon 3 of an annotated lncRNA in human and mouse genomes. The position of the MLN ORF is indicated in red.

(B) In situ hybridization showing skeletal muscle-specific expression of MLN at the indicated embryonic time points. hrt, heart; in, intestine; liv, liver; lu, lung; m, myotome; to, tongue.

(C) Northern blot of RNA isolated from adult mouse tissues using a probe specific to the full-length MLN transcript shows skeletal muscle-specific expression.

(D) Diagram of the constructs used for in vitro translation of the MLN micropeptide. The full-length MLN RNA transcript was subcloned into the CS2 vector containing the SP6 phage RNA polymerase promoter (SP6-MLN). A frameshift mutation was introduced immediately after the endogenous ATG to disrupt the MLN ORF (SP6-MLN-FS).

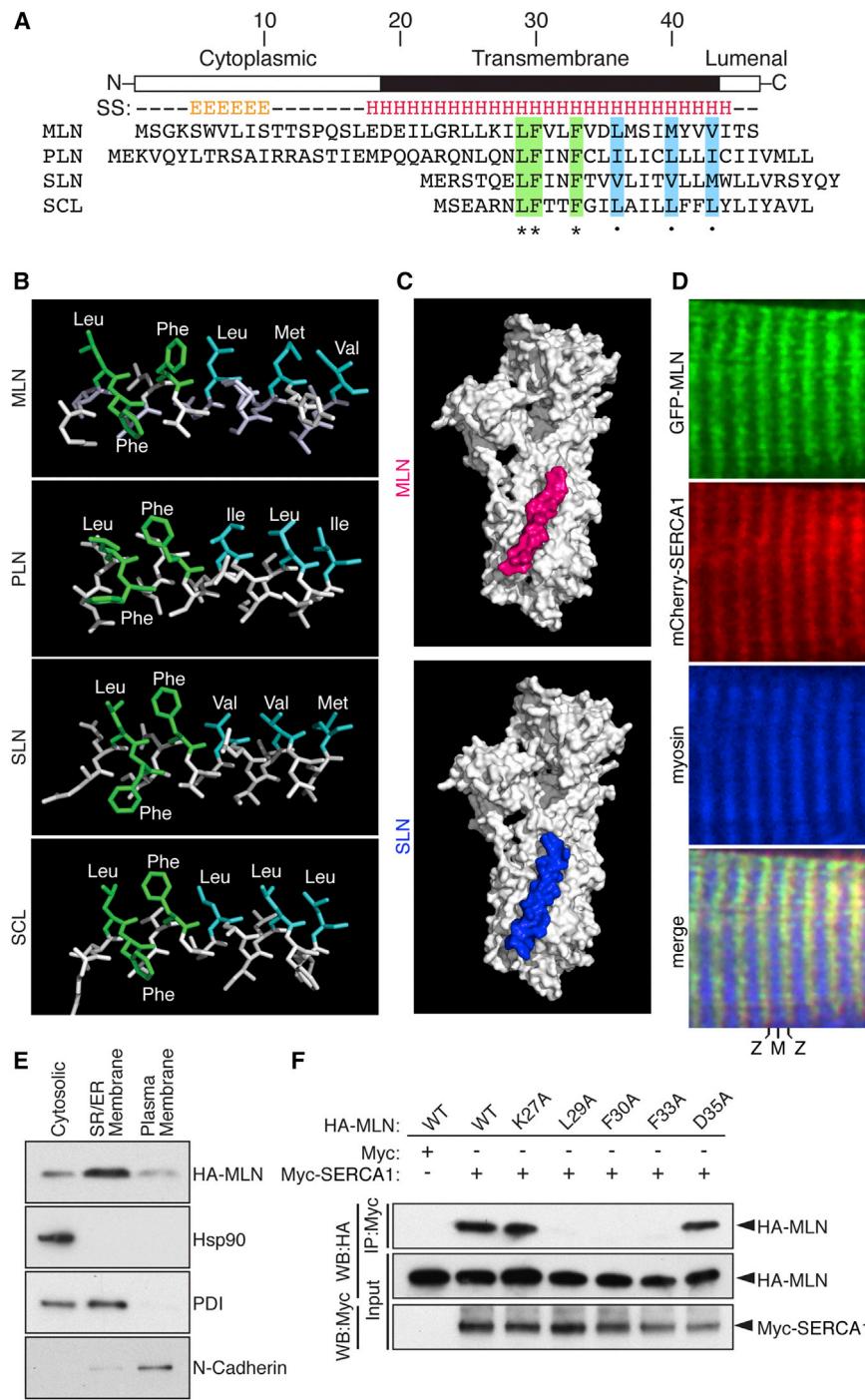
(E) Coupled in vitro transcription and translation reactions of the SP6-MLN vector using radiolabeled  $^{35}$ S-methionine produced a ~5 kDa micropeptide, visualized by Tricine SDS-PAGE. The frameshift mutation in the MLN ORF (SP6-MLN-FS) abolished any detectable expression.

(F) Targeting strategy using CRISPR/Cas9-mediated homologous recombination to knock in a FLAG epitope tag into the MLN locus in C2C12 cells. PCR-based genotyping using primers (P1–P3) or RFLP analysis of PCR products generated using primers (P4 and P5) was used to verify correct targeting.

muscles (Figures 1B and 1C). MLN transcripts are also present in C2C12 myoblasts and myotubes, but not in 10T1/2 fibroblasts (Figure S1B).

To determine if the MLN ORF is translated as a micropeptide, we transcribed and translated the full-length MLN RNA in vitro in the presence of radiolabeled  $^{35}$ S-methionine (Figures 1D and 1E). A single ~5 kDa micropeptide was produced from the MLN transcript, whereas a frameshift mutation that disrupted the MLN ORF abolished any detectable expression (Figures 1D and 1E). We further cloned a FLAG epitope tag in-frame with the C terminus of the MLN coding sequence within the full-length MLN transcript (Figure S1C). Expression of this construct in COS7 cells yielded a peptide of ~6 kDa, corresponding to the predicted molecular weight of the MLN-FLAG fusion peptide, detected by western blot (Figure S1D).

To determine whether MLN is endogenously expressed in skeletal muscle, we introduced the same FLAG epitope tag into the MLN locus in C2C12 muscle cells using CRISPR/Cas9-mediated homologous recombination (Figures 1F and S1E). PCR-based genotyping and restriction fragment-length polymorphism (RFLP) analyses were used to verify correct targeting (Figures 1G and S1F). As shown by western blot analysis in Figure 1H, C2C12 cells heterozygous for the MLN-FLAG



knockin allele expressed a micropeptide of the predicted molecular weight of the MLN-FLAG fusion peptide. Together, these data reveal that the MLN RNA, which is annotated as noncoding, in fact encodes a cryptic micropeptide.

## MLN Encodes a Transmembrane Alpha Helix

The MLN micropeptide is highly conserved across mammals and is predicted to form a type II single-pass transmembrane

**Figure 2. MLN Forms a Transmembrane Alpha Helix that Interacts with SERCA**

(A) The secondary structure (SS) of MLN is predicted to contain an N-terminal beta sheet and C-terminal transmembrane alpha helix (E, beta sheet; H, alpha helix). Alignment of mouse MLN, PLN, SLN, and the invertebrate ortholog SCL identified identical (green) and similar (blue) residues conserved in all four micropeptides.

(B) The structure of the transmembrane helices of MLN, SLN, PLN, and SCL was modeled using I-TASSER and showed a common spatial orientation of conserved residues.

(C) Automated protein docking using ClusPro predicted that MLN occupies the same groove in SERCA1 that is recognized by SLN.

(D) Two-photon laser scanning confocal microscopy showing colocalization of GFP-MLN and mCherry-SERCA1 fusion proteins in the SR of mature skeletal muscle fibers. M, M-line; Z, Z-line. (E) Retroviral expression of an N-terminal HA-tagged MLN fusion peptide (HA-MLN) in C2C12 myoblasts was enriched in the subcellular fraction containing SR/ER membrane proteins. Enrichment for cytosolic, SR/ER membrane, and plasma membrane proteins was verified by western blot analysis for Hsp90, PDI, and N-cadherin, respectively.

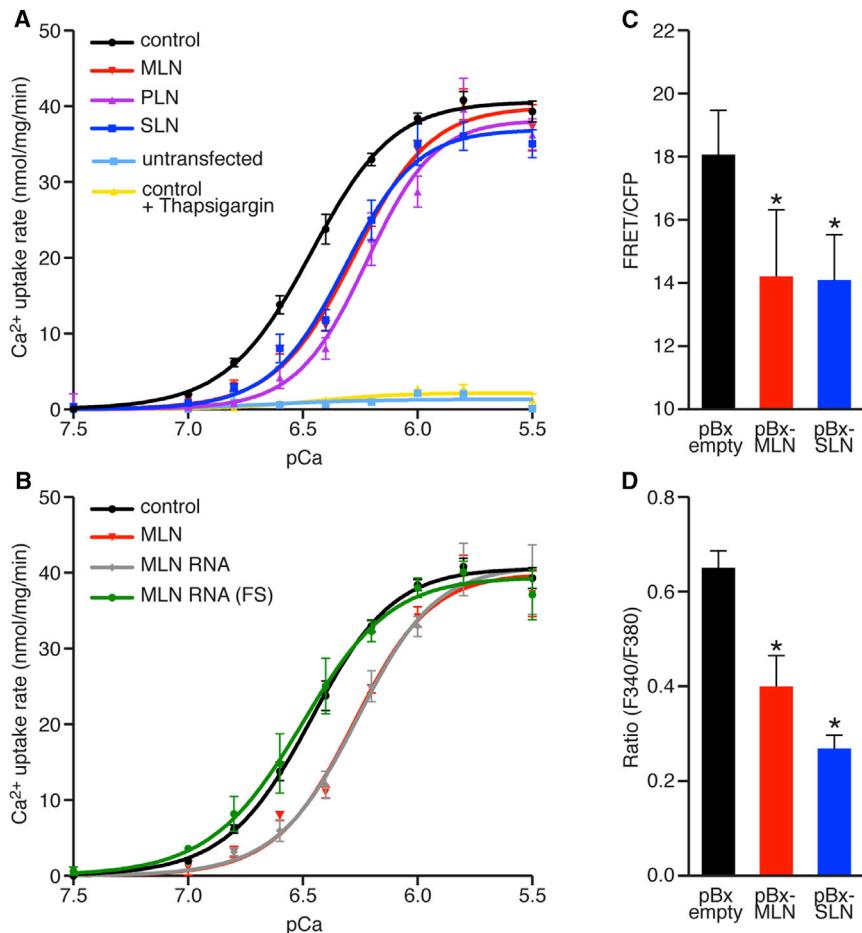
(F) Coimmunoprecipitation (coIP) experiments with HA-tagged MLN alanine mutants and a Myc-tagged SERCA1 construct transfected into COS7 cells identified residues important for interaction of MLN with SERCA1. Interaction of MLN with SERCA1 was abolished by mutation of residues shared with PLN, SLN, and SCL (L29A, F30A, and F33A), but not by the charged residues K27A or D35A. Western blot performed with anti-HA (WB-HA) or anti-Myc (WB-Myc).

See also Figure S2.

alpha helix in which the 19 N-terminal amino acid residues are exposed to the cytosol and the three C-terminal residues are luminal (Figures 2A and S2A). The N-terminal residues (5–10) of MLN are predicted to fold into a small beta sheet, and the C-terminal half of the protein contains the predicted trans-membrane alpha helix (Figures 2A and S2A). Circular dichroism spectroscopy measurements of synthesized full-length MLN peptide in detergent micelles confirmed it consists of alpha helix (21.1%), turn (9.6%), and random coil (Figure S2B).

## Association of MLN with SERCA

MLN shows a strong structural resemblance to PLN and SLN, which also form type II single-pass transmembrane peptides oriented with their N-terminal residues exposed to the cytosol. Sequence alignment of MLN, PLN, and SLN revealed a number



**Figure 3. MLN Regulates SR  $\text{Ca}^{2+}$  Levels by Inhibiting SERCA Pump Activity**

(A and B) The  $\text{Ca}^{2+}$  dependence of the relative rate of  $\text{Ca}^{2+}$  uptake is shown for homogenates from HEK293 cells cotransfected with SERCA1 and the indicated constructs. Cotransfection with MLN, PLN, or SLN resulted in a similar decrease in  $\text{Ca}^{2+}$  uptake, corresponding to a decreased affinity of SERCA for  $\text{Ca}^{2+}$ , relative to empty vector (Control). For comparison, untransfected cells and SERCA1-expressing cells treated with the SERCA inhibitor thapsigargin (100 nM) are shown. The activity of the full-length RNA transcript encoding the MLN ORF (MLN RNA) is abolished by a frameshift mutation in the MLN ORF (MLN-RNA FS).

(C) Retroviral cotransduction of C2C12 myoblasts with the FRET-based  $\text{Ca}^{2+}$  sensor T1ER with MLN or SLN was used to directly measure the relative levels of SR  $\text{Ca}^{2+}$ . Both MLN and SLN significantly decreased SR  $\text{Ca}^{2+}$  levels relative to an empty retroviral vector.

(D) Retroviral overexpression of MLN or SLN in C2C12 myoblasts treated with 4-CMC and imaged with fura-2-AM showed decreased levels of SR  $\text{Ca}^{2+}$ , measured by peak  $\text{Ca}^{2+}$  release from the SR.

Data are presented as mean  $\pm$  SEM (\* $p < 0.05$  compared to pBx-empty). See also Table S1.

A band and overlapped with the localization of an mCherry-SERCA1 fusion protein within the SR (Figure 2D). This pattern was consistent with the localization of GFP-PLN and GFP-SLN fusion proteins (Figure S2C). Additionally, an N-terminal

HA-tagged MLN fusion protein (HA-MLN) expressed in C2C12 myoblasts was enriched in the subcellular fraction containing SR/ER membrane proteins (Figure 2E), altogether suggesting that MLN functions in the membrane of the SR.

In coimmunoprecipitation experiments, the HA-MLN fusion protein formed a stable complex with SERCA1 (skeletal muscle specific), SERCA2a (cardiac and slow skeletal muscle specific), and SERCA2b (ubiquitous) isoforms (Figure S2D). Alanine mutagenesis of MLN residues shared with PLN and SLN (L29A, F30A, and F33A) abolished the ability of MLN to interact with SERCA1, whereas mutation of charged residues (K27A, D35A) did not alter this interaction (Figure 2F). These findings suggest that MLN, PLN, and SLN share a common hydrophobic binding motif that stabilizes their association with SERCA.

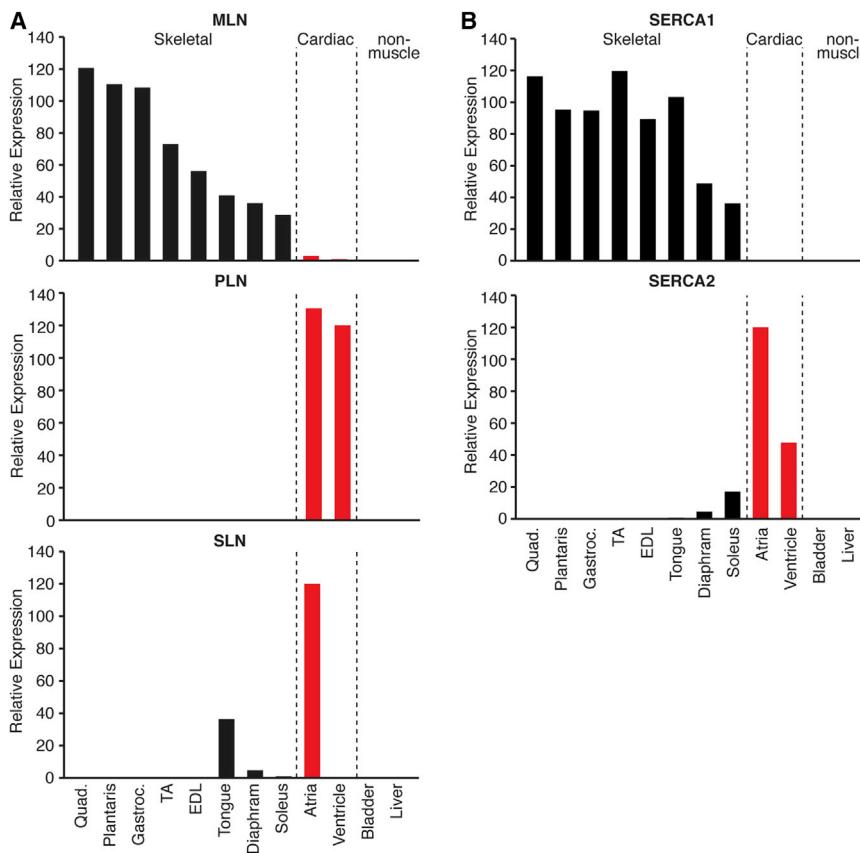
### MLN Regulates $\text{Ca}^{2+}$ Handling by Inhibiting SERCA Pump Activity

PLN and SLN both function to inhibit  $\text{Ca}^{2+}$  reuptake into the SR by lowering the affinity of SERCA for  $\text{Ca}^{2+}$ , without altering the maximal rate of  $\text{Ca}^{2+}$  pump activity ( $V_{\max}$ ). To determine if MLN regulates SERCA activity in a similar manner, we directly measured  $\text{Ca}^{2+}$ -dependent  $\text{Ca}^{2+}$ -ATPase activity in homogenates from HEK293 cells expressing SERCA1 (Figure 3A). Similar to the effects of PLN and SLN, expression of

of identically conserved residues in their transmembrane regions, which are also found in the invertebrate ortholog, sarcolamban (SCL) (Figure 2A) (Magny et al., 2013). Structural modeling of the transmembrane helices showed that all four peptides are predicted to form alpha helical structures with residues arranged in a similar spatial pattern (Figure 2B). PLN and SLN both interact with SERCA in the membrane of the SR in a common groove formed by the M2, M6, and M9 helices of SERCA (Figure 2C) (Toyoshima et al., 2003, 2013; Winther et al., 2013). Automated protein docking of the transmembrane model of MLN with the crystal structure of SERCA revealed that MLN aligned in the same groove that is occupied by PLN and SLN (Figure 2C).

### MLN Is Embedded in the SR Membrane and Colocalizes with SERCA1

To examine the subcellular localization of MLN in vivo, we electroporated a plasmid encoding GFP fused to MLN (GFP-MLN) into the flexor digitorum brevis muscle of adult mice. Two-photon laser scanning confocal microscopy allowed for the simultaneous detection of GFP fluorescence and second harmonic generation (SHG) to visualize the myosin A band of the sarcomeres (Nelson et al., 2013). The GFP-MLN fusion protein localized in a repeating pattern that alternated with the myosin



**Figure 4. Developmental and Adult Expression of MLN, PLN, and SLN in the Mouse**

(A and B) Real-time PCR showing the relative expression of MLN, PLN, SLN, and SERCA isoforms across multiple skeletal muscles, cardiac, and nonmuscle tissues isolated and pooled from three adult 8-week-old C57Bl/6 male mice. See also Figure S3.

findings demonstrate that MLN inhibits SERCA pump kinetics similar to PLN and SLN.

### MLN, PLN, and SLN Are Expressed in Different Striated Muscle Types

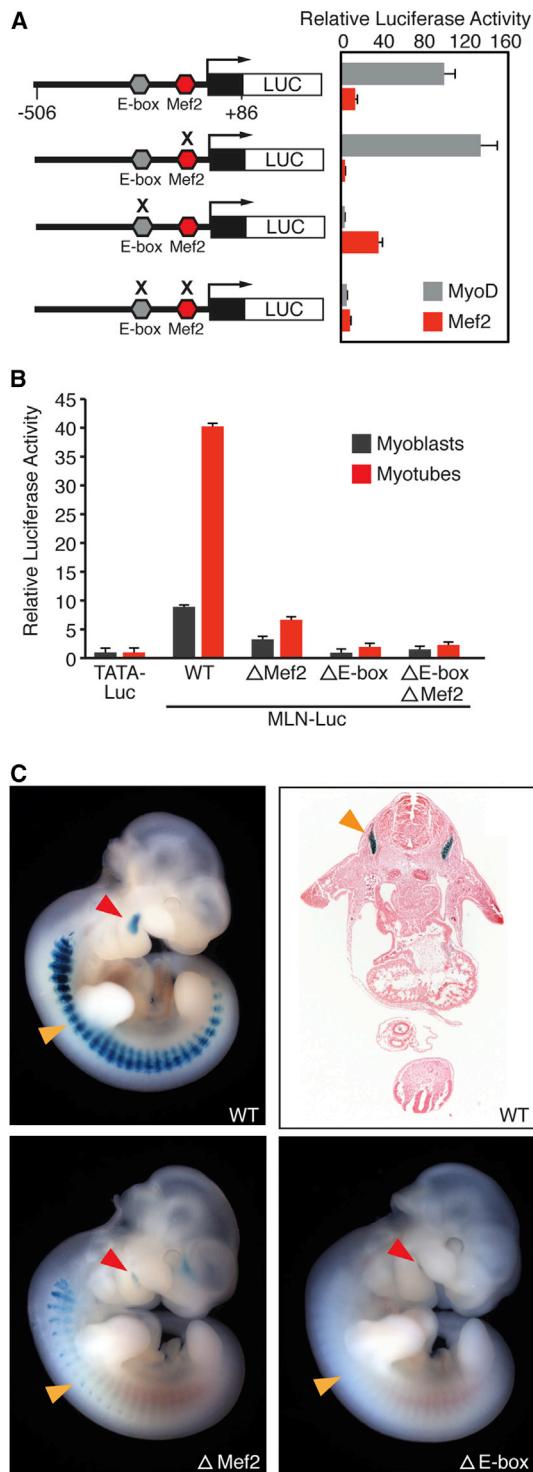
The structural and functional similarities between MLN, PLN, and SLN suggest they comprise a family of micropeptides that regulate  $\text{Ca}^{2+}$  handling through modulation of SERCA activity. To determine if PLN or SLN are functionally redundant with MLN in vivo, we examined their expression during developmental and adult stages. In the heart, PLN expression was detectable in both the atria and ventricles, whereas SLN expression was specific to the atria (Figures 4A and S3A). SLN overlapped with the expression of MLN in fetal skeletal muscles, which display a slow phenotype during embryonic development (Figure S3A) (Lu et al., 1999). Consistent with previous reports, quantitative real-time PCR and RNA sequencing (RNA-seq) expression analyses revealed that SLN was markedly downregulated in most adult skeletal muscles, which convert to fast type in the mouse (Figures 4A and S3B) (Tupling et al., 2011). In contrast, MLN was robustly expressed in all adult skeletal muscles, similar to the expression pattern of SERCA1 (Figures 4A and 4B). The expression patterns of PLN and SLN were more similar to the expression of SERCA2, which is highly expressed in cardiac and slow skeletal muscles (Figure 4B). In addition, comparison of RNA-seq expression data revealed that PLN and SLN are not expressed in differentiated C2C12 myotubes, whereas MLN transcripts are robustly expressed (Figure S3C). Thus, MLN, PLN, and SLN are differentially expressed across vertebrate muscle types, and MLN is the most abundant of the three micropeptides expressed in adult skeletal muscle of the mouse.

### Transcriptional Control of MLN Expression by MyoD and MEF2

MLN and the skeletal muscle-specific isoforms of SERCA and RyR are coregulated by MyoD, suggesting they comprise a core genetic module important for  $\text{Ca}^{2+}$  handling in skeletal muscle (Fong et al., 2012). Analysis of the 5' flanking region of the MLN gene revealed highly conserved binding sites for the myogenic transcription factors MyoD (E-box) and MEF2 (Figure S4A), which bound specifically to these sequences in gel

MLN caused a significant reduction in the rate of  $\text{Ca}^{2+}$  uptake, measured as an increase in  $K_{\text{Ca}}$  (Figure 3A; Table S1). Coexpression of the full-length RNA encoding MLN (MLN RNA) resulted in a similar decrease in  $\text{Ca}^{2+}$  uptake compared to a vector containing only the MLN coding sequence (Figure 3B). The activity of the MLN RNA was abolished by introduction of a frameshift mutation to disrupt the expression of the MLN micropeptide (MLN-RNA-FS), demonstrating that the RNA itself does not inhibit SERCA activity (Figure 3B). No effects on  $V_{\text{max}}$  were observed under any conditions tested, and the addition of thapsigargin, a potent inhibitor of SERCA activity, abolished  $\text{Ca}^{2+}$  uptake (Figure 3A).

Because the concentration of SR  $\text{Ca}^{2+}$  is dependent upon SERCA reuptake activity, we examined if MLN overexpression could alter SR  $\text{Ca}^{2+}$  levels in C2C12 myoblasts. SR  $\text{Ca}^{2+}$  levels were directly measured using retroviral delivery of the fluorescence resonance energy transfer (FRET)-based, SR-localized  $\text{Ca}^{2+}$  sensor T1ER, using culture conditions previously shown to be sensitive to changes in endogenous SERCA activity (Brandman et al., 2007). Coexpression of MLN or SLN significantly decreased the levels of SR  $\text{Ca}^{2+}$ , measured as a decrease in T1ER FRET (Figure 3C). We alternatively compared SR  $\text{Ca}^{2+}$  levels using the ratiometric  $\text{Ca}^{2+}$  indicator dye fura-2-AM, following SR  $\text{Ca}^{2+}$  release using the RyR agonist 4-chlorom-cresol (4-CMC) (Figure 3D). Consistent with our previous findings, overexpression of MLN or SLN significantly decreased peak  $\text{Ca}^{2+}$  release from the SR (Figure 3D). Together, these



**Figure 5. Regulation of MLN Transcription by MyoD and MEF2**

(A) A fragment of the MLN promoter ( $-506$  to  $+86$ , relative to the transcriptional start site) containing a highly conserved MyoD E-box (CACCTG) and MEF2 site (CTAACACAG) was cloned in front of the luciferase reporter gene (MLN-Luc). The MLN-Luc reporter was robustly transactivated by the skeletal muscle transcription factors MyoD and MEF2 in COS7 cells (gray and red bars, respectively). Mutation of the E-box (acCCgt) or Mef2 site (CTggAACAG)

mobility shift assays (Figure S4B). MyoD and MEF2 independently activated a luciferase reporter linked to the proximal MLN promoter (MLN-Luc), and transactivation was lost upon mutagenesis of their respective binding sites (Figure 5A). Chro-matin immunoprecipitation sequencing (ChIP-seq) data at this locus revealed that MyoD bound specifically to this region in both C2C12 myoblasts and myotubes (Figure S4C) (Bernstein et al., 2012). The activity of the MLN-Luc promoter in C2C12 myoblasts and myotubes was similarly dependent upon both the MEF2 and E-box sequences, which, when mutated, resulted in the loss and abrogation of MLN promoter activation, respectively (Figure 5B). Consistent with this, a LacZ transgene controlled by the MLN promoter (MLN-lacZ) displayed skeletal muscle-specific expression in vivo and was dependent upon the MEF2 and E-box sequences for full promoter activation (Figure 5C). Thus, the MLN gene is a direct target of the core transcription factors that activate skeletal myogenesis.

#### Generation of MLN Knockout Mice Using TALENs

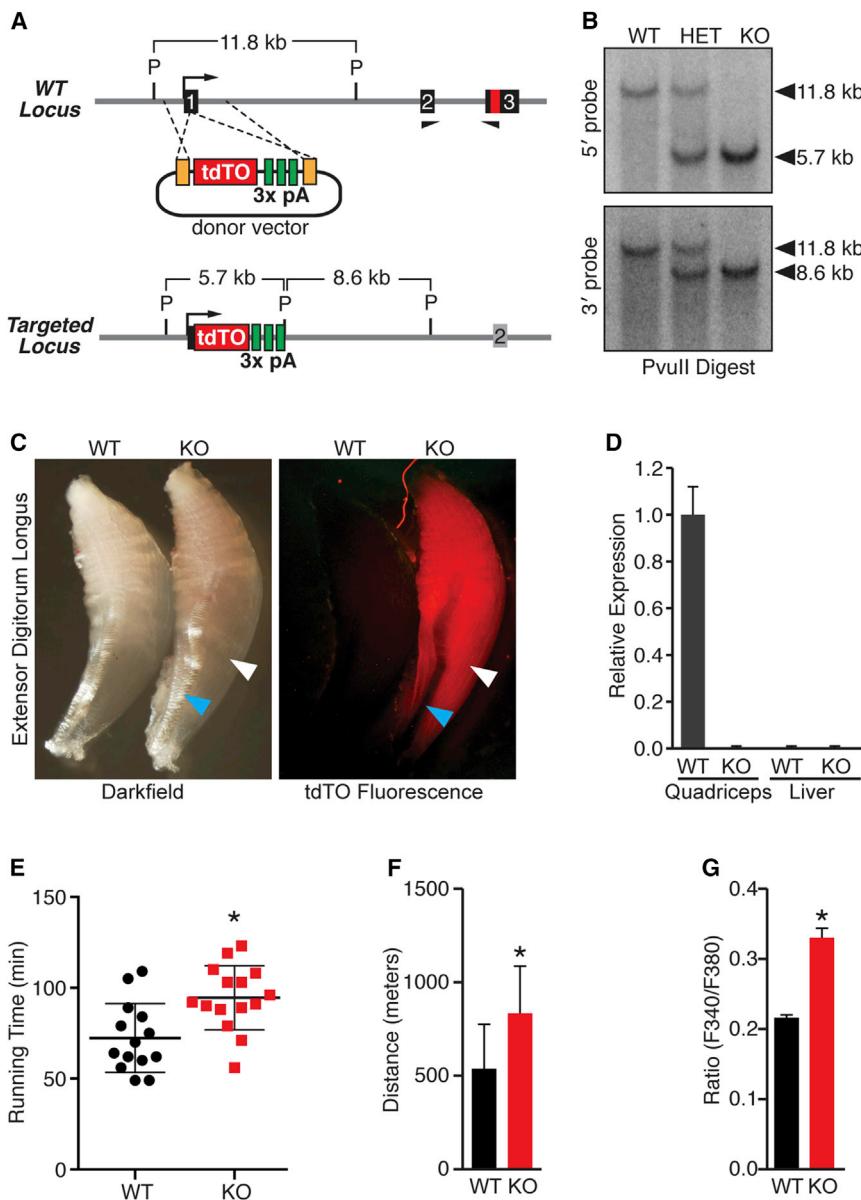
To investigate the function of MLN in vivo, we generated MLN knockout (KO) mice using TAL effector nuclease (TALEN)-mediated homologous recombination. A unique TALEN pair specific for exon 1 of the MLN locus was designed and constructed using the REAL assembly method (Reyon et al., 2012). A loss-of-function allele was created using a donor plasmid to insert a red fluorescent reporter (tdTomato), followed by a triple polyadenylation cassette into exon 1 in the MLN locus (Figure 6A). This strategy was designed to prematurely terminate transcription upstream of the MLN coding sequence but allow for the expression of a red fluorescent reporter by the endogenous MLN promoter. Correct targeting was verified by Southern blot and PCR-based genotyping (Figures 6B and S5A). Detection of tdTomato fluorescence in MLN KO mice was specific to skeletal muscle and was not detected in other tissues (Figure 6C). qPCR using primer pairs specific to the downstream exons 2 and 3 demonstrated that the MLN transcript was absent in skeletal muscle from MLN KO mice (Figure 6D).

MLN KO mice were born at expected Mendelian ratios from heterozygous intercrosses and showed no obvious morphological abnormalities or differences in body or muscle weights (Figure S5B). Mice lacking PLN and SLN also have nonpathological phenotypes but show enhanced  $\text{Ca}^{2+}$  handling and contractility in cardiac and slow skeletal muscle (Luo et al., 1994; MacLennan and Kranias, 2003; Tupling et al., 2011).

(indicated by an X) abrogated transactivation by MyoD:E12 heterodimer or Mef2c, respectively. All luciferase values were normalized to the transactivation of a basal luciferase reporter (TATA-Luc) with MyoD or Mef2, respectively. Data are presented as mean  $\pm$  SEM.

(B) Luciferase reporter assays showing that MEF2 and E-box enhancer binding sites are essential for transactivation of the MLN promoter in C2C12 myoblasts and myotubes. Data are presented as mean  $\pm$  SEM.

(C) X-gal and H&E staining of E10.5 mouse embryos harboring either the MLN promoter-lacZ transgene (WT) or mutations in the MLN promoter ( $\Delta$ Mef2 or  $\Delta$ E-box). The MLN promoter showed expression in the myotomal compartment of the somites (orange arrow) and premyogenic cells in the mandibular arch (red arrow). Mutation of the MEF2 or E-box sequences in the MLN promoter-lacZ transgene abrogated or abolished muscle-specific expression. See also Figure S4.



**Figure 6. Generation and Characterization of MLN Knockout Mice**

(A) TALEN-mediated homologous recombination was used to insert a tdTomato fluorescent reporter and triple polyadenylation cassette into exon 1 of the MLN locus to generate a null allele. A schematic of the donor vector and targeting strategy is shown. P, PvuII.

(B) Southern blot analysis confirming correct targeting of the tdTO-triple polyadenylation cassette into exon 1 of the MLN locus using probes 5' and 3' to the TALEN cut site.

(C) tdTomato fluorescence was specific to skeletal muscle (white arrowhead) of MLN KO mice and not detected in other tissues, such as tendon (blue arrowhead).

(D) Real-time PCR using primers specific to exons 2 and 3 demonstrating absence of MLN transcripts in MLN KO muscle, as well as in liver.

(E) Muscle performance was measured using forced treadmill running to exhaustion. MLN KO ( $n = 15$ ) mice ran ~31% longer than WT littermates ( $n = 14$ ).

(F) Comparison of distance run by MLN KO and WT mice in Figure 6E.

(G) Myoblasts isolated and cultured from MLN KO hindlimb muscles were imaged using Fura-2-AM and treated with the RyR agonist 4-CMC in the absence of extracellular  $\text{Ca}^{2+}$  to indirectly measure SR  $\text{Ca}^{2+}$  levels. MLN KO myoblasts showed significantly increased SR  $\text{Ca}^{2+}$  levels, measured as peak  $\text{Ca}^{2+}$  release from the SR. Data are presented as mean  $\pm$  SEM (\* $p < 0.05$  compared to WT as in E and F and pBx-empty in G).

See also Figure S5.

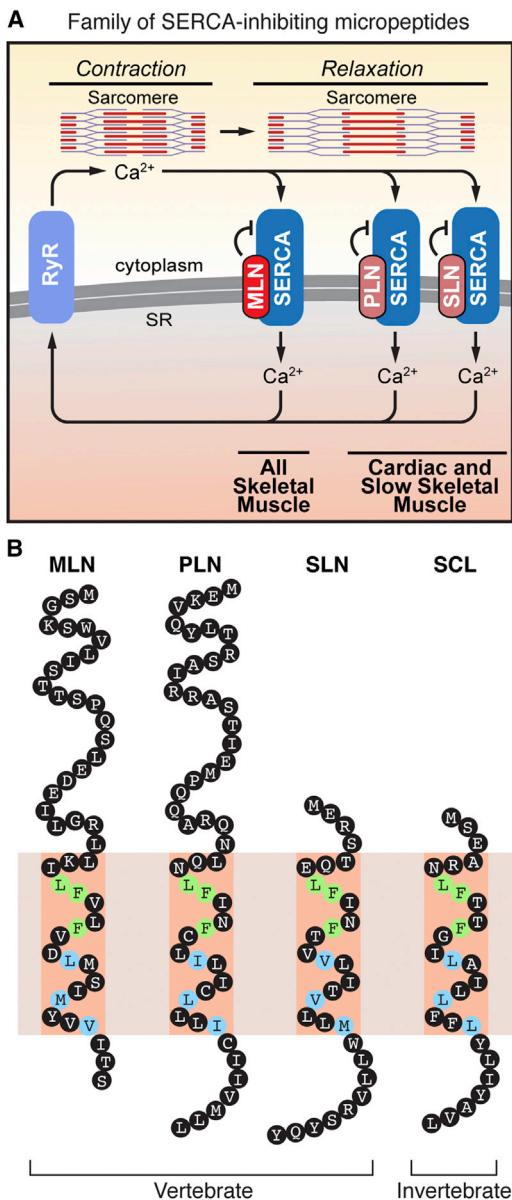
### Enhanced $\text{Ca}^{2+}$ Handling and Skeletal Muscle Performance in MLN KO Mice

To examine the potential role of MLN in regulating skeletal muscle performance, we subjected 8-week-old wild-type (WT) and MLN KO mice to a regimen of forced treadmill running to exhaustion. Remarkably, MLN KO mice ran an average time of ~31% longer than their WT littermates, representing a 55% increase in running distance (Figures 6E and 6F). Mice with a fast-to-slow fiber type switch also show increased running performance (van Rooij et al., 2009). However, histological analyses of hindlimb muscles, including quadriceps (fast-type) and soleus (slow-type) muscles revealed no obvious differences in fiber-type identity or myofiber size between WT and MLN KO mice (Figures S5C–S5E), indicating that MLN functions through a different mechanism to regulate muscle performance.

Next, we investigated whether SR  $\text{Ca}^{2+}$  levels were altered in primary myoblasts isolated from hindlimb muscles of MLN KO versus WT littermates. SR  $\text{Ca}^{2+}$  levels were measured using the calcium indicator fura-2-AM and SR  $\text{Ca}^{2+}$  release by the addition of the RyR agonist 4-CMC. Strikingly, MLN KO myoblasts showed significantly increased SR  $\text{Ca}^{2+}$  levels compared to WT myoblasts, measured as peak SR  $\text{Ca}^{2+}$  release (Figure 6G). This increase occurred without changes in the expression of RyR1 and SERCA1 expression between WT and MLN KO muscles (Figure S5F), consistent with the function of MLN as an inhibitor of SERCA activity. Together, these findings support the function of MLN as the dominant regulator of SERCA activity in adult skeletal muscle.

### DISCUSSION

SERCA is a key regulator of striated muscle performance by serving as the major  $\text{Ca}^{2+}$  ATPase responsible for the reuptake of cytosolic  $\text{Ca}^{2+}$  into the SR (Figure 7A). Direct modulation of SERCA pump activity by the micropeptides PLN and SLN regulates muscle contractility by diminishing the rate of  $\text{Ca}^{2+}$  reuptake into the SR. However, because PLN and SLN are not



**Figure 7. A Family of SERCA-Inhibitory Micropeptides**

(A) RyR and SERCA play a critical role in muscle contractility by controlling  $\text{Ca}^{2+}$  cycling between the cytosol and SR. MLN, PLN, and SLN inhibit SERCA pump activity in different striated muscle types of vertebrates.

(B) Illustration of the family of SERCA-inhibitory micropeptides. The discovery of MLN reveals that vertebrates encode three SERCA inhibitory peptides that share conserved residues within their transmembrane alpha helices. Green shading denotes identical residues, and blue shading denotes similar residues. MLN, myoregulin; PLN, phospholamban; SLN, sarcolipin; SCL, sarcolamban.

expressed in most adult skeletal muscles of mice, the possible existence of other SERCA regulatory factors has remained an open question.

Our biochemical and *in vivo* results show that MLN forms a stable complex with SERCA in the membrane of the SR and that MLN directly influences SR  $\text{Ca}^{2+}$  levels and maximal

exercise performance. The robust skeletal muscle-specific expression of MLN, resulting from direct transactivation by the myogenic transcription factors MyoD and MEF2, further highlight MLN as the predominant SERCA regulating micropeptide in adult skeletal muscle. The discovery of MLN reveals a universal mechanism for the control of SERCA activity by a family of related micropeptides expressed in different striated muscle types in vertebrates (Figure 7A).

### Micropeptides Can Be Concealed within RNAs

#### Misannotated as Noncoding

Micropeptides remain highly underrepresented in genome annotations due in large part to the difficulty in identifying functional short ORFs in RNA transcripts. Recent advances in bioinformatic and biochemical methodologies have revealed that lncRNAs may harbor concealed micropeptides; however, only a few have been functionally verified and characterized *in vivo*. Here, we have identified and characterized the function of a conserved micropeptide in vertebrates that functions as an important regulator of skeletal muscle performance through modulation of  $\text{Ca}^{2+}$  handling by SERCA. Given the attention that has been focused on the control of SERCA activity and muscle function over the years, it is remarkable that such a key regulator of these processes has gone undetected. Undoubtedly, this is because the MLN ORF is concealed in an RNA annotated as noncoding. Our discovery of MLN was possible due to its high sequence conservation in vertebrates and the presence of an identifiable functional domain, a type II transmembrane alpha helix. We found that the MLN alpha helix shares residues in common with PLN and SLN; however, common bioinformatic search tools alone are not sufficient to identify the relatedness of MLN, PLN, and SLN. Interestingly, both PLN and SLN were first discovered as micropeptides and subsequently mapped to RNA transcripts (Kirchberger et al., 1975; Wawrzynow et al., 1992). The PLN and SLN genes share a genomic architecture similar to that of MLN, in which the small ORFs are encoded in the 5' region of the terminal exon. Indeed, the transcripts encoding PLN and SLN may have also been annotated as noncoding, if not for their prior discovery as micropeptides.

### A Family of SERCA-Inhibitory Micropeptides

In *Drosophila*, the invertebrate ortholog of SERCA is encoded by a single gene (Ca-P60A) and is modulated by the recently identified micropeptide sarcolamban (SCL) (Magny et al., 2013). In vertebrates, the SERCA family has expanded to encode three genes (SERCA1-3) that give rise to multiple alternate splice variants with differing kinetic properties and expression patterns (Anger et al., 1994; Periasamy and Kalyanasundaram, 2007). The expansion of vertebrate gene families occurred through whole-genome and gene-duplication events, resulting in the formation of paralogous gene families. Combined with the tissue-specific expression patterns of the individual SERCA isoforms, the differential expression of MLN, PLN, and SLN likely contributes to the unique  $\text{Ca}^{2+}$  handling and contractile properties of different striated muscle types in vertebrates. In addition, the extent to which MLN, PLN, and SLN partially overlap in their expression in different muscle types may influence the calcium kinetics and performance of these tissues. Coexpression of

PLN and SLN in atrial cardiomyocytes and slow skeletal muscle has been shown to result in the superinhibition of SERCA pump activity (MacLennan et al., 2003). MLN expression overlaps with that of PLN and SLN in adult slow skeletal muscle and with SLN in developing skeletal muscles. Future biochemical and animal studies are required to determine the extent to which MLN synergizes with PLN and SLN in coregulating SERCA activity in slow-type and developing skeletal muscle. SLN expression in larger animals is more widespread than in rodents, occurring in both fast- and slow-type muscles (Fajardo et al., 2013; Odermatt et al., 1997). A synergistic interaction between MLN and SLN may be more biologically relevant in the adult tissues of these species; however, the relative expression patterns of these two genes in these species remain to be determined.

Apart from spatiotemporal differences in expression, the SERCA-regulatory peptides differ in size and presence of additional secondary structures. A schematic illustration of the family of SERCA-inhibitory micropeptides is shown in Figure 7B. Both SLN and SCL lack extended N-terminal regions and additional secondary structures other than their transmembrane alpha helices. MLN and PLN have expanded N-terminal cytoplasmic regions that encode a beta sheet and alpha helix, respectively. The N-terminal sequence of PLN has been shown to be critical for its function, and phosphorylation of serine-16 by protein kinase A (PKA) or threonine-17 by  $\text{Ca}^{2+}$ /calmodulin-dependent protein kinase II (CaMKII) diminishes the ability of PLN to inhibit SERCA activity (Mattiazzi et al., 2006; Wegener et al., 1989). The N-terminal sequence of MLN could serve as a similar regulatory domain through phosphorylation, as this sequence contains multiple serine and threonine residues. Deciphering the physiological signaling pathways that regulate MLN expression and function will be important to fully understand its role in skeletal muscle development and disease.

### Future Questions

Defects in  $\text{Ca}^{2+}$  signaling underlie the pathogenesis of many muscle diseases that arise from mutations in components of  $\text{Ca}^{2+}$  signaling pathways, as well as diseases that arise from a loss of myofiber structural integrity (Berchtold et al., 2000; Millay et al., 2008). Given the importance of SERCA pump activity in regulating  $\text{Ca}^{2+}$  handling and the pathogenesis of skeletal muscle diseases, such as Brody myopathy and muscular dystrophies (Allen et al., 2010; Goonasekera et al., 2011; Odermatt et al., 1996), the discovery of MLN opens interesting possibilities for the modulation of these pathways. Considering the enhanced exercise capacity of MLN KO mice, it is interesting to speculate that pharmacologic approaches to disrupt the association of MLN with SERCA might have similar salutary effects. Modulation of SERCA activity in skeletal muscle has also been implicated in the control of systemic energy homeostasis (Bal et al., 2012), raising the interesting possibility that MLN may exert additional metabolic functions. Considering that SERCA is also important in  $\text{Ca}^{2+}$  regulation in nonmuscle cell types in which MLN, PLN, and SLN are not expressed, it is interesting to speculate that additional SERCA-modulating micropeptides may be concealed within putative lncRNAs expressed in other tissues. Finally, the discovery of MLN as a previously unrecognized regulator of muscle function suggests that the microproteome, which is

largely unexplored, represents a reservoir for future biological insights.

### EXPERIMENTAL PROCEDURES

#### TALEN-Mediated Homologous Recombination in Mice

A unique TALEN pair specific for the MLN locus was designed using the ZifIT Targeter Program (<http://zifit.partners.org/ZifIT/Introduction.aspx>) and constructed using the REAL Assembly Kit (Addgene) (Reyon et al., 2012). A donor vector containing the tdTomato reporter and triple polyadenylation sequences was constructed by incorporating short 5' and 3' homology arms specific to the MLN locus. TALEN mRNAs were in vitro transcribed using the mMessage mMachine T7 Ultra Kit (Life Technologies), diluted to  $25 \text{ ng } \mu\text{l}^{-1}$ , and coinjected with  $3 \text{ ng } \mu\text{l}^{-1}$  of the circular DNA donor plasmid into the nucleus and cytoplasm of one-cell-stage embryos (B6C3F1) and transferred into pseudopregnant female ICR mice.

#### Study Approval

All experimental procedures involving animals in this study were reviewed and approved by the University of Texas Southwestern Medical Center's Institutional Animal Care and Use Committee.

#### Radioisotopic In Situ Hybridization

In situ hybridizations were performed as previously described (Shelton et al., 2000). See the Extended Experimental Procedures for a more detailed protocol. The primer sequences used to clone MLN, PLN, and SLN cDNA templates are listed in Table S2.

#### CRISPR/Cas9-Mediated Homologous Recombination in C2C12 Myoblasts

A single-guide RNA (sgRNA) specific to the C-terminal coding sequence of the mouse MLN locus was cloned into the sgRNA/Cas9 expression vector px330 (MLN-FLAG-px330). The donor vector was constructed with a single FLAG epitope tag in-frame with the MLN coding sequence flanked by ~500 base pair homology arms specific to the MLN locus. MLN-FLAG knockin clones were generated by transient cotransfection and expanded from single-cell clones. Detection of endogenous MLN-FLAG peptide was performed by immunoblotting with a rabbit anti-FLAG antibody (Sigma) on protein lysates immunoprecipitated with mouse anti-FLAG agarose beads (Sigma).

#### Treadmill Running

In blinded studies, male MLN KO and WT littermate mice were subjected to forced exercise on a treadmill (Exer-6M, Columbus Instruments, 10% incline) at 8 weeks of age using a regimen previously described (van Rooij et al., 2009). See the Extended Experimental Procedures for a more detailed protocol.

#### Intracellular $\text{Ca}^{2+}$ Imaging

Cytosolic  $\text{Ca}^{2+}$  levels were measured as described previously (Liou et al., 2005), with the exception that retroviral-transduced C2C12 cells or primary myoblasts were plated on Ibidi  $\mu$ -35 mm tissue-culture dishes and cultured for 24 hr in low  $\text{Ca}^{2+}$  ( $0.1 \text{ mM Ca}^{2+}$ ) prior to imaging to increase sensitivity to changes in SR  $\text{Ca}^{2+}$  levels (Brandman et al., 2007). SR  $\text{Ca}^{2+}$  levels were directly measured using T1ER as previously described (Abell et al., 2011; Tsai et al., 2014). See the Extended Experimental Procedures for a more detailed protocol.

#### Oxalate-Supported $\text{Ca}^{2+}$ Uptake Measurements in HEK293 Lysates

Oxalate-supported  $\text{Ca}^{2+}$ -dependent  $\text{Ca}^{2+}$ -ATPase activity in homogenates was measured by a modification of the Millipore filtration technique as described previously (Holemans et al., 2014; Luo et al., 1994). HEK293 cells were cotransfected with equal amounts of an expression plasmid encoding mouse SERCA1 and an expression plasmid encoding MLN, PLN, or SLN. See the Extended Experimental Procedures for a more detailed protocol.

#### Coimmunoprecipitations and Western Blot Analysis

Coimmunoprecipitations (coIPs) were performed as previously described (Anderson et al., 2009). Tris-tricine-SDS-PAGE was carried out using 16.5%

Tris-tricine gels (BioRad) and Tris-tricine-SDS running buffer (BioRad). See the [Extended Experimental Procedures](#) for a more detailed protocol.

#### Real-Time PCR

Total RNA was prepared from whole muscles using Trizol (Invitrogen) and treated with DNase prior to reverse transcription by Superscript III (Invitrogen). Real-time PCR was performed using TaqMan probes (ABI) or SYBR green using primers in [Table S2](#). TaqMan probes include Mck (Mm00432556\_m1), Met2c (Mm01340842\_m1), SLN (Mm00481536\_m1), PLN (Mm00452263\_m1), and SERCA2 (Mm01201434\_m1). Primers for SYBR green reactions are listed in [Table S2](#).

#### Electrophoretic Mobility Shift Assays

Electrophoretic mobility shift assays (EMSA) were performed as previously described ([Anderson et al., 2012](#)) using recombinant myc-tagged proteins expressed in COS7 cells and double-stranded EMSA probes created by annealing complementary oligonucleotides. Probe sequences are listed in [Table S2](#). Supershifts were performed by adding 1 µg of mouse anti-Myc antibody (Invitrogen). See the [Extended Experimental Procedures](#) for a more detailed protocol.

#### Luciferase Assays

Luciferase assays were performed as previously described ([Anderson et al., 2009](#)). Luciferase activity was measured using a FluoStar OPTIMA microplate reader (BMG Labtech) and normalized to beta-galactosidase activity using the FluoReporter LacZ/Galactosidase Quantitation Kit (Invitrogen). See the [Extended Experimental Procedures](#) for a more detailed protocol.

#### Subcellular Fractionation

C2C12 myoblasts infected with a retrovirus encoding the HA-MLN fusion protein (pBabeX-HA-MLN) were fractionated as previously described ([Millay et al., 2013](#)). See the [Extended Experimental Procedures](#) for a more detailed protocol.

#### Northern and Southern Blot Analysis

Northern blots were performed using a commercially prepared adult mouse multitissue RNA blot (MN-MT-1; ZymoGen) hybridized with a radiolabeled DNA probe specific to the full-length MLN transcript. Radiolabeled DNA probes for northern and Southern blots were generated using a RadPrime Kit (Invitrogen) ([Table S2](#)).

#### Histology and Immunohistochemistry

Skeletal muscle tissues were dissected and fixed overnight in 4% formaldehyde in PBS prior to paraffin embedding and sectioning using routine procedures. Immunohistochemistry was performed on deparaffinized sections using a HistoMouse-Plus kit (Invitrogen) using primary antibodies specific to fast (MY32; Sigma) and slow (NOQ7.54; Sigma) myosins. Wheat germ agglutinin (WGA) staining was performed using Alexa Fluor 555-conjugated WGA (Invitrogen) as described previously ([Liu et al., 2011](#)).

#### Analysis of RNA-Seq Expression Data

Raw data for C2C12 cells and triceps brachii muscle were downloaded from the Short Read Archive (SRP002119 and SRP008123, respectively). Reads were mapped to the UCSC mm9 genome annotation using TopHat, and alignments were processed to bigWig coverage maps and viewed using the UCSC genome browser.

#### Circular Dichroism Spectroscopy

Circular dichroism (CD) spectroscopy measurements were performed using a JASCO J-815 spectrometer on *in vitro*-synthesized full-length MLN (Peptide 2.0). The secondary structure elements were calculated using Yang's fit. The root-mean-square deviation for the observed and calculated CD spectra values was 3%. See the [Extended Experimental Procedures](#) for a more detailed protocol.

#### Structural Modeling and Automated Protein Docking

The helical domains of MLN, PLN, SLN, and SCL were created ab initio using I-TASSER ([Zhang, 2008](#)). Automated protein docking of the MLN model with

the crystal structure of SERCA1 (4H1W) was performed using ClusPro 2.0 ([Comeau et al., 2004](#)).

#### Adult Muscle Electroporation and Imaging

Flexor digitorum brevis (FDB) muscles of 12-week-old male mice were electroporated as previously described ([Nelson et al., 2013](#)), with expression vectors encoding N-terminal GFP fusions to MLN, PLN, or SLN. Unfixed FDB muscles were examined directly using two-photon laser scanning microscopy (Zeiss; LSM 780), with reverse second harmonic generation to visualize the A bands as an internal reference. See the [Extended Experimental Procedures](#) for a more detailed protocol.

#### SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, five figures, and two tables and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2015.01.009>.

#### ACKNOWLEDGMENTS

We thank Dr. Tobias Meyer at Stanford University Medical Center for generously providing the pcDNA-T1ER vector and Dr. David H. MacLennan at University of Toronto for generously providing the Serca1 (A52) antibody. This work was supported by grants from the NIH (HL-077439, HL-111665, HL-093039, DK-099653, and U01-HL-100401), Fondation Leducq Networks of Excellence, and the Robert A. Welch Foundation (grant 1-0025 to E.N.O.; I-1789 to J.L.). D.M.A. was supported by an American Heart Association postdoctoral fellowship (13POST14570050). K.M.A. was supported by an American Heart Association predoctoral fellowship (14PRE19830031). B.R.N. was supported by an NIH Training grant (1F30AR067094-01), and P.K. was supported by a postdoctoral grant from the Sigrid Juselius Foundation.

Received: August 19, 2014

Revised: November 24, 2014

Accepted: January 5, 2015

Published: January 29, 2015

#### REFERENCES

- Abell, E., Ahrends, R., Bandara, S., Park, B.O., and Teruel, M.N. (2011). Parallel adaptive feedback enhances reliability of the Ca<sup>2+</sup> signaling system. Proc. Natl. Acad. Sci. USA 108, 14485–14490.
- Allen, D.G., Gervasio, O.L., Yeung, E.W., and Whitehead, N.P. (2010). Calcium and the damage pathways in muscular dystrophy. Can. J. Physiol. Pharmacol. 88, 83–91.
- Anderson, D.M., Beres, B.J., Wilson-Rawls, J., and Rawls, A. (2009). The homeobox gene Mohawk represses transcription by recruiting the sin3A/HDAC co-repressor complex. Dev. Dyn. 238, 572–580.
- Anderson, D.M., George, R., Noyes, M.B., Rowton, M., Liu, W., Jiang, R., Wolfe, S.A., Wilson-Rawls, J., and Rawls, A. (2012). Characterization of the DNA-binding properties of the Mohawk homeobox transcription factor. J. Biol. Chem. 287, 35351–35359.
- Andrews, S.J., and Rothnagel, J.A. (2014). Emerging evidence for functional peptides encoded by short open reading frames. Nat. Rev. Genet. 15, 193–204.
- Anger, M., Samuel, J.L., Marotte, F., Wuytack, F., Rappaport, L., and Lompré, A.M. (1994). In situ mRNA distribution of sarco(endo)plasmic reticulum Ca(2+)-ATPase isoforms during ontogeny in the rat. J. Mol. Cell. Cardiol. 26, 539–550.
- Bal, N.C., Maurya, S.K., Sopariwala, D.H., Sahoo, S.K., Gupta, S.C., Shaikh, S.A., Pant, M., Rowland, L.A., Bombardier, E., Goonasekera, S.A., et al. (2012). Sarcolipin is a newly identified regulator of muscle-based thermogenesis in mammals. Nat. Med. 18, 1575–1579.
- Bassel-Duby, R., and Olson, E.N. (2006). Signaling pathways in skeletal muscle remodeling. Annu. Rev. Biochem. 75, 19–37.

Bazzini, A.A., Johnstone, T.G., Christiano, R., Mackowiak, S.D., Obermayer, B., Fleming, E.S., Vejnar, C.E., Lee, M.T., Rajewsky, N., Walther, T.C., and Giraldez, A.J. (2014). Identification of small ORFs in vertebrates using ribosome footprinting and evolutionary conservation. *EMBO J.* 33, 981–993.

Berchtold, M.W., Brinkmeier, H., and Müntener, M. (2000). Calcium ion in skeletal muscle: its crucial role for muscle function, plasticity, and disease. *Physiol. Rev.* 80, 1215–1265.

Bernstein, B.E., Birney, E., Dunham, I., Green, E.D., Gunter, C., and Snyder, M.; ENCODE Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature* 489, 57–74.

Berridge, M.J., Bootman, M.D., and Roderick, H.L. (2003). Calcium signalling: dynamics, homeostasis and remodelling. *Nat. Rev. Mol. Cell Biol.* 4, 517–529.

Brandman, O., Liou, J., Park, W.S., and Meyer, T. (2007). STIM2 is a feedback regulator that stabilizes basal cytosolic and endoplasmic reticulum Ca<sup>2+</sup> levels. *Cell* 131, 1327–1339.

Briggs, F.N., Lee, K.F., Wechsler, A.W., and Jones, L.R. (1992). Phospholamban expressed in slow-twitch and chronically stimulated fast-twitch muscles minimally affects calcium affinity of sarcoplasmic reticulum Ca(2+)-ATPase. *J. Biol. Chem.* 267, 26056–26061.

Chu, G., Ferguson, D.G., Edes, I., Kiss, E., Sato, Y., and Kranias, E.G. (1998). Phospholamban ablation and compensatory responses in the mammalian heart. *Ann. N Y Acad. Sci.* 853, 49–62.

Comeau, S.R., Gatchell, D.W., Vajda, S., and Camacho, C.J. (2004). ClusPro: an automated docking and discrimination method for the prediction of protein complexes. *Bioinformatics* 20, 45–50.

Dorn, G.W., 2nd, and Molkentin, J.D. (2004). Manipulating cardiac contractility in heart failure: data from mice and men. *Circulation* 109, 150–158.

Fajardo, V.A., Bombardier, E., Vigna, C., Devji, T., Bloomberg, D., Gamu, D., Gramolini, A.O., Quadrilatero, J., and Tupling, A.R. (2013). Co-expression of SERCA isoforms, phospholamban and sarcolipin in human skeletal muscle fibers. *PLoS ONE* 8, e84304.

Fong, A.P., Yao, Z., Zhong, J.W., Cao, Y., Ruzzo, W.L., Gentleman, R.C., and Tapscott, S.J. (2012). Genetic and epigenetic determinants of neurogenesis and myogenesis. *Dev. Cell* 22, 721–735.

Goonasekera, S.A., Lam, C.K., Millay, D.P., Sargent, M.A., Hajjar, R.J., Kranias, E.G., and Molkentin, J.D. (2011). Mitigation of muscular dystrophy in mice by SERCA overexpression in skeletal muscle. *J. Clin. Invest.* 121, 1044–1052.

Holemans, T., Vandecaetsbeek, I., Wuytack, F., and Vangheluwe, P. (2014). Measuring Ca<sup>2+</sup>-dependent Ca<sup>2+</sup>-uptake activity in the mouse heart. *Cold Spring Harb Protoc* 2014, 876–886.

Kirchberger, M.A., Tada, M., and Katz, A.M. (1975). Phospholamban: a regulatory protein of the cardiac sarcoplasmic reticulum. *Recent Adv. Stud. Cardiac Struct. Metab.* 5, 103–115.

Kranias, E.G., and Hajjar, R.J. (2012). Modulation of cardiac contractility by the phospholamban/SERCA2a regulatome. *Circ. Res.* 110, 1646–1660.

Liou, J., Kim, M.L., Heo, W.D., Jones, J.T., Myers, J.W., Ferrell, J.E., Jr., and Meyer, T. (2005). STIM is a Ca<sup>2+</sup> sensor essential for Ca<sup>2+</sup>-store-depletion-triggered Ca<sup>2+</sup> influx. *Curr. Biol.* 15, 1235–1241.

Liu, N., Bezprozvannaya, S., Shelton, J.M., Frisard, M.I., Hulver, M.W., McMillan, R.P., Wu, Y., Voelker, K.A., Grange, R.W., Richardson, J.A., et al. (2011). Mice lacking microRNA 133a develop dynamin 2-dependent centronuclear myopathy. *J. Clin. Invest.* 121, 3258–3268.

Lu, B.D., Allen, D.L., Leinwand, L.A., and Lyons, G.E. (1999). Spatial and temporal changes in myosin heavy chain gene expression in skeletal muscle development. *Dev. Biol.* 216, 312–326.

Luo, W., Grupp, I.L., Harrer, J., Ponniah, S., Grupp, G., Duffy, J.J., Doetschman, T., and Kranias, E.G. (1994). Targeted ablation of the phospholamban gene is associated with markedly enhanced myocardial contractility and loss of beta-agonist stimulation. *Circ. Res.* 75, 401–409.

MacLennan, D.H., and Kranias, E.G. (2003). Phospholamban: a crucial regulator of cardiac contractility. *Nat. Rev. Mol. Cell Biol.* 4, 566–577.

MacLennan, D.H., Asahi, M., and Tupling, A.R. (2003). The regulation of SERCA-type pumps by phospholamban and sarcolipin. *Ann. N Y Acad. Sci.* 986, 472–480.

Magny, E.G., Pueyo, J.I., Pearl, F.M., Cespedes, M.A., Niven, J.E., Bishop, S.A., and Couso, J.P. (2013). Conserved regulation of cardiac calcium uptake by peptides encoded in small open reading frames. *Science* 341, 1116–1120.

Mattiazzi, A., Mundina-Weilenmann, C., Vittone, L., Said, M., and Kranias, E.G. (2006). The importance of the Thr17 residue of phospholamban as a phosphorylation site under physiological and pathological conditions. *Braz. J. Med. Biol. Res.* 39, 563–572.

Millay, D.P., Sargent, M.A., Osinska, H., Baines, C.P., Barton, E.R., Vuagniaux, G., Sweeney, H.L., Robbins, J., and Molkentin, J.D. (2008). Genetic and pharmacologic inhibition of mitochondrial-dependent necrosis attenuates muscular dystrophy. *Nat. Med.* 14, 442–447.

Millay, D.P., O'Rourke, J.R., Sutherland, L.B., Bezprozvannaya, S., Shelton, J.M., Bassel-Duby, R., and Olson, E.N. (2013). Myomaker is a membrane activator of myoblast fusion and muscle formation. *Nature* 499, 301–305.

Minamisawa, S., Wang, Y., Chen, J., Ishikawa, Y., Chien, K.R., and Matsuoka, R. (2003). Atrial chamber-specific expression of sarcolipin is regulated during development and hypertrophic remodeling. *J. Biol. Chem.* 278, 9570–9575.

Nelson, B.R., Wu, F., Liu, Y., Anderson, D.M., McAnally, J., Lin, W., Cannon, S.C., Bassel-Duby, R., and Olson, E.N. (2013). Skeletal muscle-specific T-tubule protein STAC3 mediates voltage-induced Ca<sup>2+</sup> release and contractility. *Proc. Natl. Acad. Sci. USA* 110, 11881–11886.

Odermatt, A., Taschner, P.E., Khanna, V.K., Busch, H.F., Karpati, G., Jablecki, C.K., Breuning, M.H., and MacLennan, D.H. (1996). Mutations in the gene-encoding SERCA1, the fast-twitch skeletal muscle sarcoplasmic reticulum Ca<sup>2+</sup>-ATPase, are associated with Brody disease. *Nat. Genet.* 14, 191–194.

Odermatt, A., Taschner, P.E., Scherer, S.W., Beatty, B., Khanna, V.K., Cornblath, D.R., Chaudhry, V., Yee, W.C., Schrank, B., Karpati, G., et al. (1997). Characterization of the gene encoding human sarcolipin (SLN), a proteolipid associated with SERCA1: absence of structural mutations in five patients with Brody disease. *Genomics* 45, 541–553.

Pan, Y., Zvaritch, E., Tupling, A.R., Rice, W.J., de Leon, S., Rudnicki, M., McKerlie, C., Banwell, B.L., and MacLennan, D.H. (2003). Targeted disruption of the ATP2A1 gene encoding the sarco(endo)plasmic reticulum Ca<sup>2+</sup>-ATPase isoform 1 (SERCA1) impairs diaphragm function and is lethal in neonatal mice. *J. Biol. Chem.* 278, 13367–13375.

Periasamy, M., and Kalyanasundaram, A. (2007). SERCA pump isoforms: their role in calcium transport and disease. *Muscle Nerve* 35, 430–442.

Reyon, D., Khayter, C., Regan, M.R., Joung, J.K., and Sander, J.D. (2012). Engineering designer transcription activator-like effector nucleases (TALENs) by REAL or REAL-Fast assembly. *Curr. Protoc. Mol. Biol. Chapter 12, Unit 12. 15.*

Rossi, A.E., and Dirksen, R.T. (2006). Sarcoplasmic reticulum: the dynamic calcium governor of muscle. *Muscle Nerve* 33, 715–731.

Schmitt, J.P., Kamisago, M., Asahi, M., Li, G.H., Ahmad, F., Mende, U., Kranias, E.G., MacLennan, D.H., Seidman, J.G., and Seidman, C.E. (2003). Dilated cardiomyopathy and heart failure caused by a mutation in phospholamban. *Science* 299, 1410–1413.

Shelton, J.M., Lee, M.H., Richardson, J.A., and Patel, S.B. (2000). Microsomal triglyceride transfer protein expression during mouse development. *J. Lipid Res.* 41, 532–537.

Slack, J.P., Grupp, I.L., Luo, W., and Kranias, E.G. (1997). Phospholamban ablation enhances relaxation in the murine soleus. *Am. J. Physiol.* 273, C1–C6.

Tada, M., and Toyofuku, T. (1998). Molecular regulation of phospholamban function and expression. *Trends Cardiovasc. Med.* 8, 330–340.

Toyoshima, C., Asahi, M., Sugita, Y., Khanna, R., Tsuda, T., and MacLennan, D.H. (2003). Modeling of the inhibitory interaction of phospholamban with the Ca<sup>2+</sup> ATPase. *Proc. Natl. Acad. Sci. USA* 100, 467–472.

Toyoshima, C., Iwasawa, S., Ogawa, H., Hirata, A., Tsueda, J., and Inesi, G. (2013). Crystal structures of the calcium pump and sarcolipin in the Mg<sup>2+</sup>-bound E1 state. *Nature* 495, 260–264.

Tsai, F.C., Seki, A., Yang, H.W., Hayer, A., Carrasco, S., Malmersjö, S., and Meyer, T. (2014). A polarized Ca<sup>2+</sup>, diacylglycerol and STIM1 signalling system regulates directed cell migration. *Nat. Cell Biol.* 16, 133–144.

Tupling, A.R., Bombardier, E., Gupta, S.C., Hussain, D., Vigna, C., Bloemberg, D., Quadrilatero, J., Trivieri, M.G., Babu, G.J., Backx, P.H., et al. (2011). Enhanced Ca<sup>2+</sup> transport and muscle relaxation in skeletal muscle from sarcolipin-null mice. *Am. J. Physiol. Cell Physiol.* 301, C841–C849.

van Rooij, E., Quiat, D., Johnson, B.A., Sutherland, L.B., Qi, X., Richardson, J.A., Kelm, R.J., Jr., and Olson, E.N. (2009). A family of microRNAs encoded by myosin genes governs myosin expression and muscle performance. *Dev. Cell* 17, 662–673.

Vangheluwe, P., Schuermans, M., Zádor, E., Waelkens, E., Raeymaekers, L., and Wuytack, F. (2005). Sarcolipin and phospholamban mRNA and protein expression in cardiac and skeletal muscle of different species. *Biochem. J.* 389, 151–159.

Wawrzynow, A., Theibert, J.L., Murphy, C., Jona, I., Martonosi, A., and Collins, J.H. (1992). Sarcolipin, the “proteolipid” of skeletal muscle sarcoplasmic reticulum, is a unique, amphipathic, 31-residue peptide. *Arch. Biochem. Biophys.* 298, 620–623.

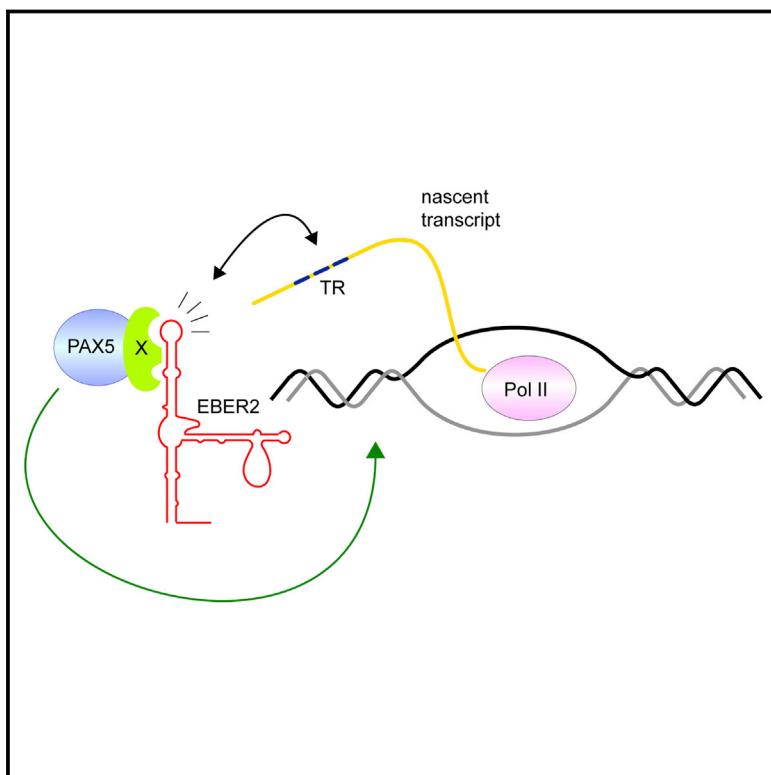
Wegener, A.D., Simmerman, H.K., Lindemann, J.P., and Jones, L.R. (1989). Phospholamban phosphorylation in intact ventricles. Phosphorylation of serine 16 and threonine 17 in response to beta-adrenergic stimulation. *J. Biol. Chem.* 264, 11468–11474.

Winther, A.M., Bublitz, M., Karlsen, J.L., Møller, J.V., Hansen, J.B., Nissen, P., and Buch-Pedersen, M.J. (2013). The sarcolipin-bound calcium pump stabilizes calcium sites exposed to the cytoplasm. *Nature* 495, 265–269.

Zhang, Y. (2008). I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* 9, 40.

# EBV Noncoding RNA Binds Nascent RNA to Drive Host PAX5 to Viral DNA

## Graphical Abstract



## Authors

Nara Lee, Walter N. Moss, Therese A. Yario, Joan A. Steitz

## Correspondence

joan.steitz@yale.edu

## In Brief

A highly abundant nuclear noncoding RNA expressed by the Epstein-Barr virus localizes to specific sites on viral chromatin and facilitates binding of the host transcription factor PAX5 through RNA-RNA interactions.

## Highlights

- A viral noncoding RNA localizes to repeat regions of the viral genome
- Viral noncoding RNA enables recruitment of a cellular transcription factor to DNA
- Recruitment requires RNA-RNA interaction between noncoding RNA and nascent RNA
- Depletion of noncoding RNA results in decreased viral lytic replication

# EBV Noncoding RNA Binds Nascent RNA to Drive Host PAX5 to Viral DNA

Nara Lee,<sup>1</sup> Walter N. Moss,<sup>1</sup> Therese A. Yario,<sup>1</sup> and Joan A. Steitz<sup>1,\*</sup>

<sup>1</sup>Department of Molecular Biophysics and Biochemistry, Howard Hughes Medical Institute, Yale University School of Medicine, 295 Congress Avenue, New Haven, CT 06536, USA

\*Correspondence: joan.steitz@yale.edu

<http://dx.doi.org/10.1016/j.cell.2015.01.015>

## SUMMARY

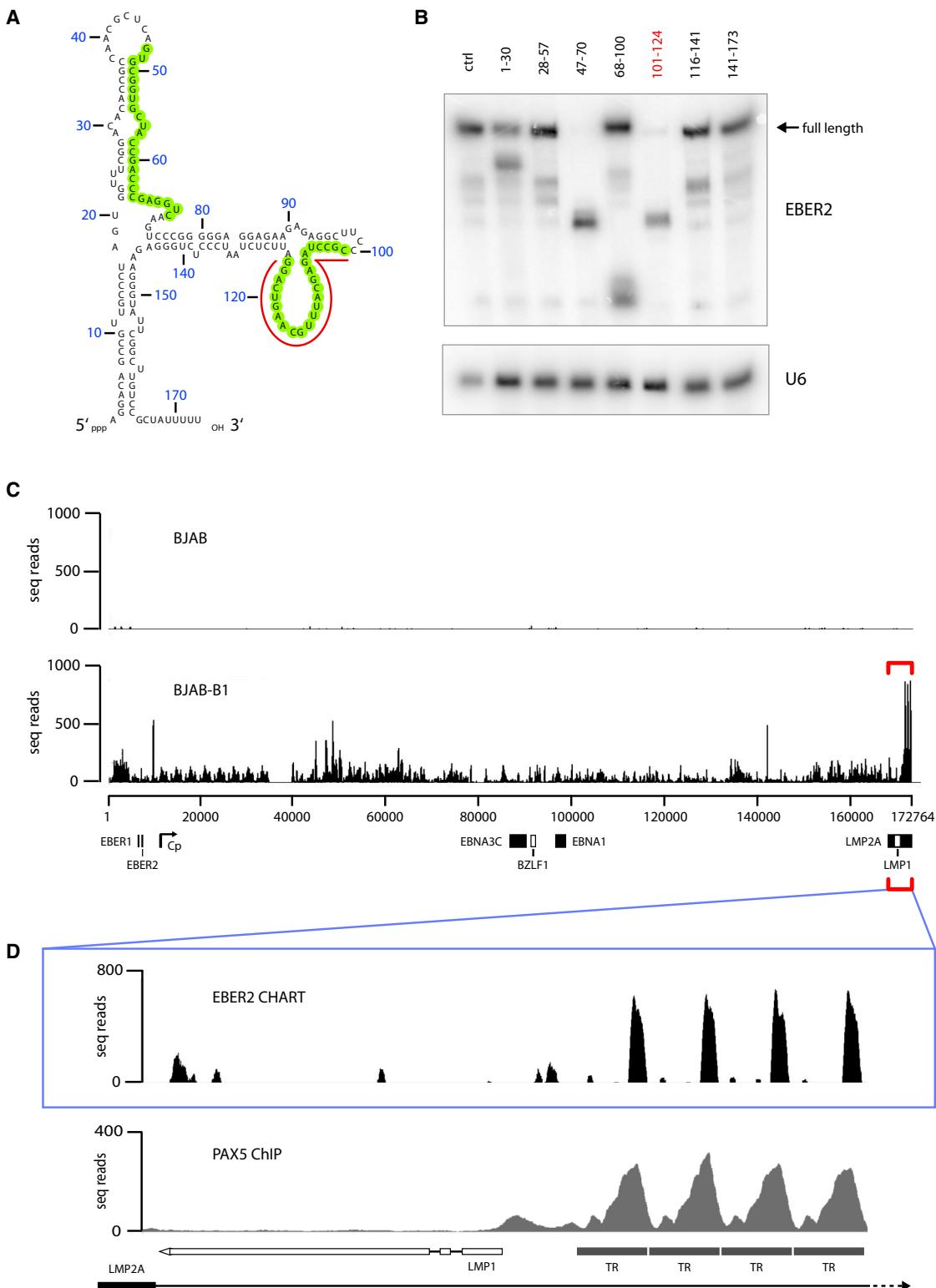
EBER2 is an abundant nuclear noncoding RNA expressed by the Epstein-Barr virus (EBV). Probing its possible chromatin localization by CHART revealed EBER2's presence at the terminal repeats (TRs) of the latent EBV genome, overlapping previously identified binding sites for the B cell transcription factor PAX5. EBER2 interacts with PAX5 and is required for the localization of PAX5 to the TRs. EBER2 knockdown phenocopies PAX5 depletion in upregulating the expression of LMP2A/B and LMP1, genes nearest the TRs. Knockdown of EBER2 also decreases EBV lytic replication, underscoring the essential role of the TRs in viral replication. Recruitment of the EBER2-PAX5 complex is mediated by base-pairing between EBER2 and nascent transcripts from the TR locus. The interaction is evolutionarily conserved in the related primate herpesvirus CeHV15 despite great sequence divergence. Using base-pairing with nascent RNA to guide an interacting transcription factor to its DNA target site is a previously undescribed function for a *trans*-acting noncoding RNA.

## INTRODUCTION

Epstein-Barr virus (EBV) is a human lymphotropic gamma-1 herpesvirus (or lymphocryptovirus, LCV) that expresses two noncoding RNAs called EBER1 (EBV-encoded RNA 1) and EBER2, which are 167 and 173 nucleotides (nts) long, respectively (Lerner et al., 1981). They are expressed during all forms of EBV latency and also during lytic growth (Greifenegger et al., 1998; Rowe et al., 2009). EBER1 accumulates to  $\sim 10^6$  and EBER2 to  $\sim 2.5 \times 10^5$  copies per infected cell (Moss and Steitz, 2013). The high copy number and the evolutionary conservation of EBERs in related primate LCVs point to a fundamental role of EBERs in the EBV life cycle (Howe and Shu, 1988; Rivailler et al., 2002b). To elucidate their function, recombinant EBV strains lacking EBERs have been engineered and introduced into host B lymphocytes. However, conflicting observations regarding possible effects on B cell growth and transformation have been reported (Gregorovic et al., 2011; Swaminathan et al., 1991; Yajima et al., 2005). Thus, the physiological roles of EBERs have remained unclear. Likewise, only limited

mechanistic insights have been gained from investigating the interacting partners of these noncoding RNAs. Both EBER1 and EBER2 bind the host RNA chaperone protein La (Lerner et al., 1981), whereas ribosomal protein L22 and AU1 (AU-rich element binding factor 1)/hnRNP D (heterogeneous nuclear ribonucleoprotein D) additionally associate with EBER1 (Lee et al., 2012; Toczyski et al., 1994). A hallmark of EBV is its tumorigenic potential, and several EBV latent proteins have been shown to contribute to oncogenicity (Young and Rickinson, 2004). Intriguingly, EBERs by themselves, particularly EBER1, can cause tumors under certain conditions, but the exact molecular mechanism has not been elucidated even though EBER1's interaction with L22 has been implicated in the process (Houmani et al., 2009; Komano et al., 1999; Repellin et al., 2010). The unidentified function(s) of EBERs must be confined to the nucleus, as they exhibit strictly nuclear localization and do not undergo nucleo-cytoplasmic shuttling (Fok et al., 2006).

The EBV genome in virions is linear, flanked on both ends by tandem terminal repeats (TRs). These direct repeat units are 538 and 544 base pairs (bps) long (in EBV type I and II strains, respectively) and contain high GC content (78%). Each viral genome contains a varying number of up to 20 TRs (Brown et al., 1988). Upon infection of a host cell, the viral genome circularizes at the TRs, possibly through a recombination event, and amplifies as multicopy episomes during latency (Lindahl et al., 1976; Sugden et al., 1979). Following genome circularization, the promoter region and exons of LMP2 (Latent Membrane Protein 2), located at opposite ends of the linear genome, become juxtaposed to allow the expression of both LMP2A and LMP2B isoforms, which differ by alternative promoter usage (Raab-Traub and Flynn, 1986). Both isoforms modulate B cell receptor signal transduction to prevent premature lytic reactivation (Miller et al., 1994; Rovedo and Longnecker, 2007). During lytic replication, the circular EBV genome is amplified, giving rise to long concatemers that are subsequently processed into unit length genomes. Again, processing occurs in the TR region by a proposed recombination event in conjunction with enzymatic cleavage by a so-called terminase complex (Chiu et al., 2014; Zimmermann and Hammerschmidt, 1995). The TRs further provide an essential sorting signal for the linear genome to be packaged into capsids (Feederle et al., 2005). These observations underscore the important role(s) of the TR regions in EBV genome organization during both latency and the lytic cycle. The latency-lytic switch is subject to tight regulation, as an expanding body of evidence indicates that lytic replication contributes to oncogenesis (Katsumura et al., 2012; Ma et al., 2011).



**Figure 1. EBER2 Localizes to the TRs of the EBV Genome**

(A) Secondary structure model of EBER2. RNase H-sensitive regions (shown in B) are indicated in green. The region hybridizing to the ASO used in CHART is underlined.

(B) Northern blot of EBER2 after RNase H digestion using DNA oligonucleotides complementary to EBER2. The numbers on top correspond to the nucleotides targeted in EBER2. An arrow indicates the mobility of full-length EBER2 RNA. U6 RNA serves as a loading control.

(legend continued on next page)

B lymphocytes are the major cell type infected by EBV. The transcription factor PAX5 (Paired box protein 5) has been described as a master regulator of B lymphocyte development through promoting the expression of B cell specific genes and repressing B-lineage inappropriate genes (Medvedovic et al., 2011). Specific DNA binding of PAX5 is achieved through the conserved paired box DNA-binding motif. Intriguingly, a recent study has shown that PAX5 binds to the TRs of EBV, perhaps to coordinate viral genome organization (Arvey et al., 2012). PAX5 has further been shown to regulate EBV latent gene expression, as depletion of PAX5 results, for example, in upregulation of the expression of LMP1, the main transforming protein of EBV that acts as a classical oncogene, and of both LMP2 isoforms (Arvey et al., 2012).

Here, we performed capture hybridization analysis of RNA targets (CHART) for EBER2, a method comparable to chromatin immunoprecipitation (ChIP), but in CHART the chromatin localization of an RNA of interest is assayed (Simon et al., 2011). In lieu of an antibody in ChIP experiments, CHART employs antisense oligonucleotides (ASOs) that hybridize to accessible regions of an RNA for selection (Figure S1). We report that EBER2 localizes to the TRs of the latent EBV genome and provide *in vivo* evidence that its recruitment involves an RNA-RNA interaction with nascent RNA transcripts. This process in turn is required for efficient association of PAX5 with its target sites within the TRs. Perturbation of EBER2-PAX5 localization affects expression of genes nearest its binding site as well as lytic viral DNA replication, with possible downstream effects on oncogenic processes.

## RESULTS

### EBER2 Co-Localizes with PAX5 to the TRs of the EBV Genome

To identify an accessible region in EBER2 that could be targeted by an ASO for selection in CHART (Simon et al., 2011), we added ~30-nt-long DNA oligonucleotides complementary to EBER2 to lysate from EBV-positive BJAB-B1 cells (which contain type II EBV). Formation of DNA-RNA hybrids at accessible regions in EBER2 induces cleavage by endogenous RNase H. Two such regions in EBER2 (nucleotides 47–70 and 101–124) were detected by Northern blot analysis (Figures 1A and 1B). To select EBER2, we therefore coupled to agarose beads an RNA ASO targeting nucleotides 101–124, as the secondary structure of this region is predicted to form an extensive loop (Figure 1A).

We then used CHART to identify EBER2 binding sites on chromatin in the EBV-positive BJAB-B1 cell line; the isogenic EBV-deficient BJAB cell line served as a negative control (Figure S1). Deep sequencing libraries from both cell lines were prepared after CHART and subjected to Illumina massive parallel sequencing. When the sequencing reads were mapped to the host cell genome, no obvious EBER2 peaks were present in infected BJAB-B1 cells compared to BJAB cells (data not shown). However, prominent EBER2 binding sites mapped to the 3' end

of the annotated EBV genome (Figure 1C, bottom, bracketed region). Since very few sequence reads from control BJAB cells map to the EBV genome (Figure 1C, top), these peaks are unlikely to represent host sequences that misalign with viral DNA. A zoomed-in view shows that EBER2 localizes to the TR regions of the EBV genome (Figure 1D, top), its profile strikingly overlapping published ChIP data for the transcription factor PAX5 (Figure 1D). Because TRs represent tandem repeat sequences, as for PAX5 (Arvey et al., 2012), we cannot distinguish whether EBER2 binds to only one specific TR or whether it is equally distributed across all TRs, as depicted here.

Given their co-localization on EBV chromatin, we asked whether EBER2 and PAX5 interact with each other. Co-immunoprecipitation after *in vivo* formaldehyde crosslinking using anti-PAX5 antibody showed that EBER2 interacts with PAX5, while EBER2 was not co-precipitated using an immunoglobulin G (IgG) control antibody (Figures 2A and S2A). A reciprocal experiment was performed using an EBER2 ASO (complementary to nucleotides 101–124) that should select EBER2-associated proteins. As shown by western blot analysis, PAX5 was enriched by the EBER2 ASO, while a control ASO against EBER1 failed to capture PAX5 (Figures 2B and S2B). We asked whether EBER2 interacts directly with PAX5 by performing an RNA immunoprecipitation (IP) assay under denaturing conditions after UV crosslinking (Lee et al., 2012). EBER2 did not precipitate with anti-PAX5 antibody (Figure S2C), consistent with the fact that EBER2 does not exhibit a band-shift in the presence of recombinant Pax5 in electrophoretic mobility shift assays (EMSA) (Figures S2D–S2G). Together, these results suggest that EBER2 and PAX5 interact, but the association may be indirect.

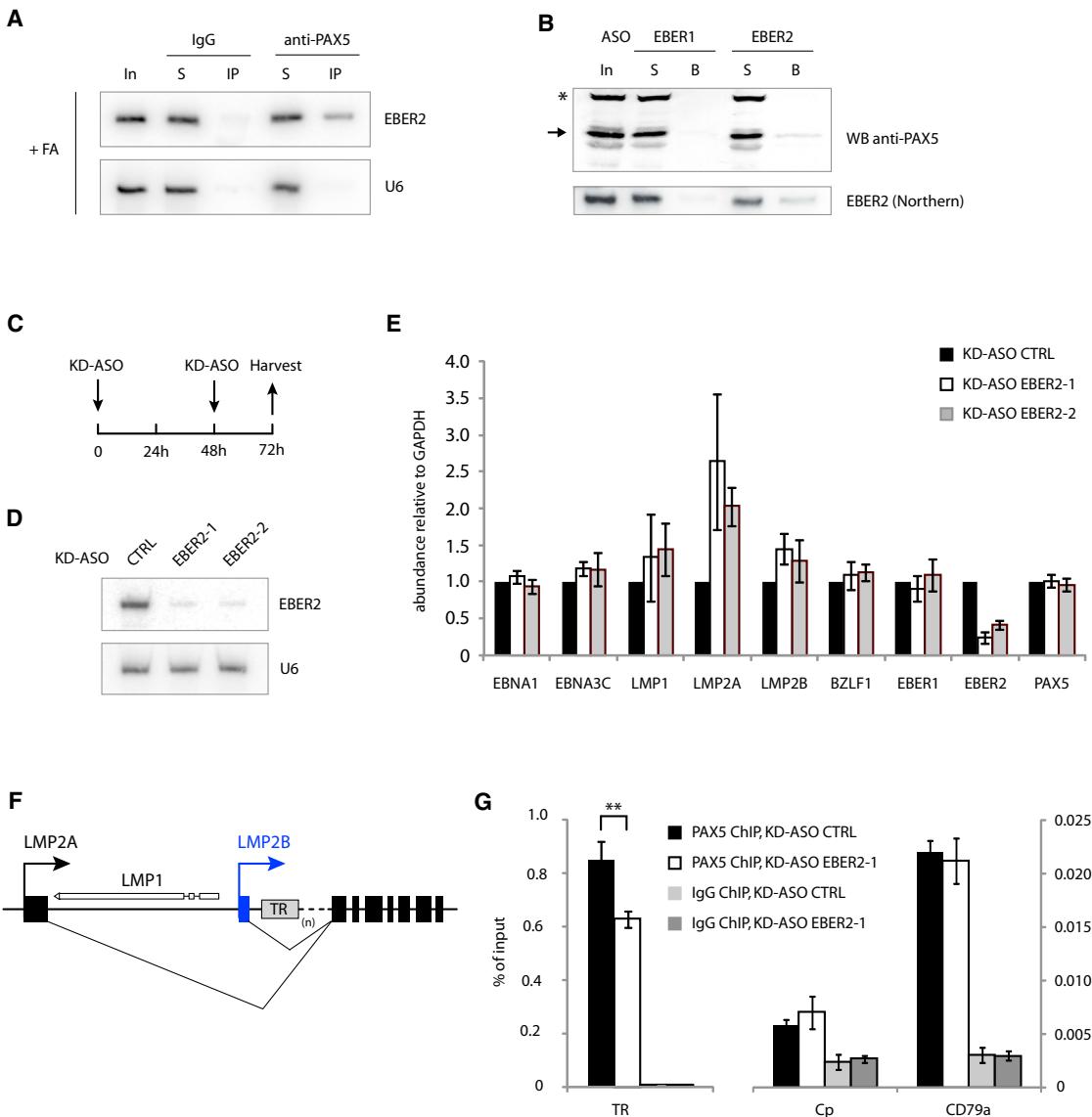
Based on its interaction with PAX5, we reasoned that EBER2 might act in concert with PAX5 to regulate EBV-latent genes. Therefore, we knocked down EBER2 using chimeric ASOs that induce endogenous RNase H-mediated degradation (Table S1) (Ideue et al., 2009) and assessed the mRNA levels of several EBV-latent genes by qRT-PCR. Two knockdown ASOs (KD ASOs) that target the available regions in EBER2 (nucleotides 101–124 and 39–62; Figures 1A and 1B) efficiently depleted EBER2 to less than 20% of its original level upon nucleofection (Figure 2D). As latent gene expression at 48 hr post-nucleofection did not change (data not shown), we introduced a second KD ASO nucleofection step at 48 hr and harvested the cells after 3 days of depletion (Figure 2C). This procedure was necessary to maintain EBER2 at less than 20% its original level because EBER2 levels increased from ~18% at 24 hr to ~46% at 48 hr after a single knockdown.

Upon EBER2 depletion, we observed that expression of LMP2A, and to a lesser extent LMP1 and LMP2B, was upregulated (Figure 2E), phenocopying the results of PAX5 knockdown (Arvey et al., 2012). We observed no significant change for other EBV genes (e.g., EBNA1 and BZLF1), as reported for PAX5 depletion (Arvey et al., 2012), possibly because we examined the RNA levels at an earlier time point (after 3 days of EBER2

(C) EBER2-CHART results from BJAB and BJAB-B1 cells. Deep sequencing reads were mapped to the entire EBV genome (x axis); the number of sequence reads is plotted on the y axis. Several EBV genes and the C promoter region are indicated. White boxes for BZLF1 and LMP1 indicate reverse gene orientation.

(D) EBER2-CHART peaks in the TR region (bracket in C) and the PAX5 ChIP profile (from Arvey et al. 2012) are shown.

See also Figure S1 and Table S1.



**Figure 2. EBER2 Interacts with PAX5 and Is Required for Efficient PAX5 Binding to TRs**

(A) Northern blot of EBER2 after IP with IgG (control) or anti-PAX5 antibody after formaldehyde crosslinking (+FA). In, 5% input; S, 5% supernatant; IP, 100%. (B) EBER1 and EBER2 ASOs were used to pull down associated proteins, followed by western blot using anti-PAX5 antibody (top). An arrow indicates PAX5; an asterisk indicates a non-specific band. In, 10% input; S, 10% supernatant; B, 100% beads. The same samples were subjected to northern blot analysis to detect EBER2 (bottom). Quantification of (A) and (B) are shown in Figures S2A and S2B.

(C) Experimental outline for EBER2 knockdown.

(D) Northern blot for EBER2 was carried out after EBER2 knockdown with two different KD ASOs (complementary to nucleotides 101–124 and 39–62) that target the nucleotides indicated in green in Figure 1A. The same blot was probed for U6 as a loading control.

(E) RNA levels of several EBV genes were assessed by qRT-PCR after EBER2 knockdown.

(F) The LMP locus of the episomal EBV genome. LMP1 is transcribed in the opposite direction to LMP2. The variable number of TRs is indicated by (n).

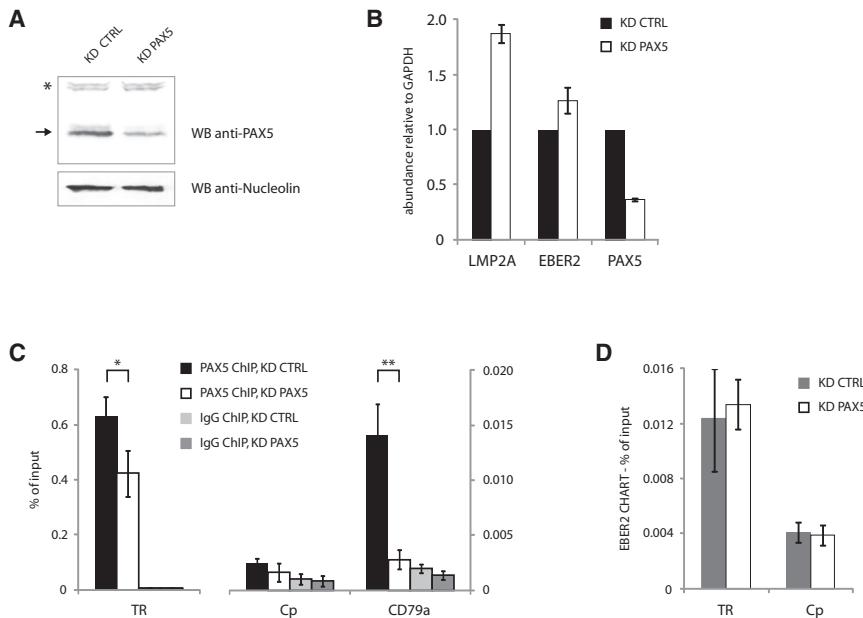
(G) PAX5 localization at the TRs after EBER2 knockdown was measured by ChIP-qPCR. The cellular CD79a promoter region, a known PAX5 target site, served as a positive ChIP control. The C promoter region of the EBV genome (Cp), an active promoter region not bound by PAX5, was the negative control. All data represent the mean of three independent experiments  $\pm$  SD; \*\*p = 0.008 (Student's t test; n = 3).

See also Figure S2 and Table S1.

depletion as compared to after 5 days of PAX5 depletion) in an effort to reduce potential secondary effects of knockdown. Since the TRs are located in the first intron of LMP2A/B, positioned close to the transcription start site of LMP1 (Figure 2F), it is not

unexpected for these genes to be most affected by lack of EBER2 localization to the TRs.

Finally, we asked whether EBER2 knockdown would affect PAX5 recruitment to the TRs. As shown by PAX5 ChIP, EBER2



**Figure 3. PAX5 Is Dispensable for EBER2 Recruitment to the TRs**

(A) Knockdown efficiency of PAX5 by lentivirally expressed shRNA was determined by western blot using anti-PAX5 antibody. Anti-Nucleolin antibody provided a loading control.

(B) qRT-PCR analysis after PAX5 knockdown.

(C) PAX5 ChIP-qPCR analysis after PAX5 knockdown. \* $p = 0.03$ , \*\* $p = 0.002$  (Student's *t* test;  $n = 3$ ). (D) EBER2-CHART followed by qPCR analysis after PAX5 knockdown. All data represent the mean of three independent experiments  $\pm$  SD.

See also Table S1.

depletion specifically reduced PAX5 localization at the TRs, whereas its binding was unaltered at a cellular PAX5 target, the CD79a promoter region (Figure 3G) (Revilla-I-Domingo et al., 2012). In summary, our results show that EBER2 is required for PAX5 recruitment to the TRs and synergizes with PAX5 to regulate certain EBV latent genes.

#### Base Pairing of EBER2 to Nascent LMP2 Transcripts Recruits PAX5 to the TRs

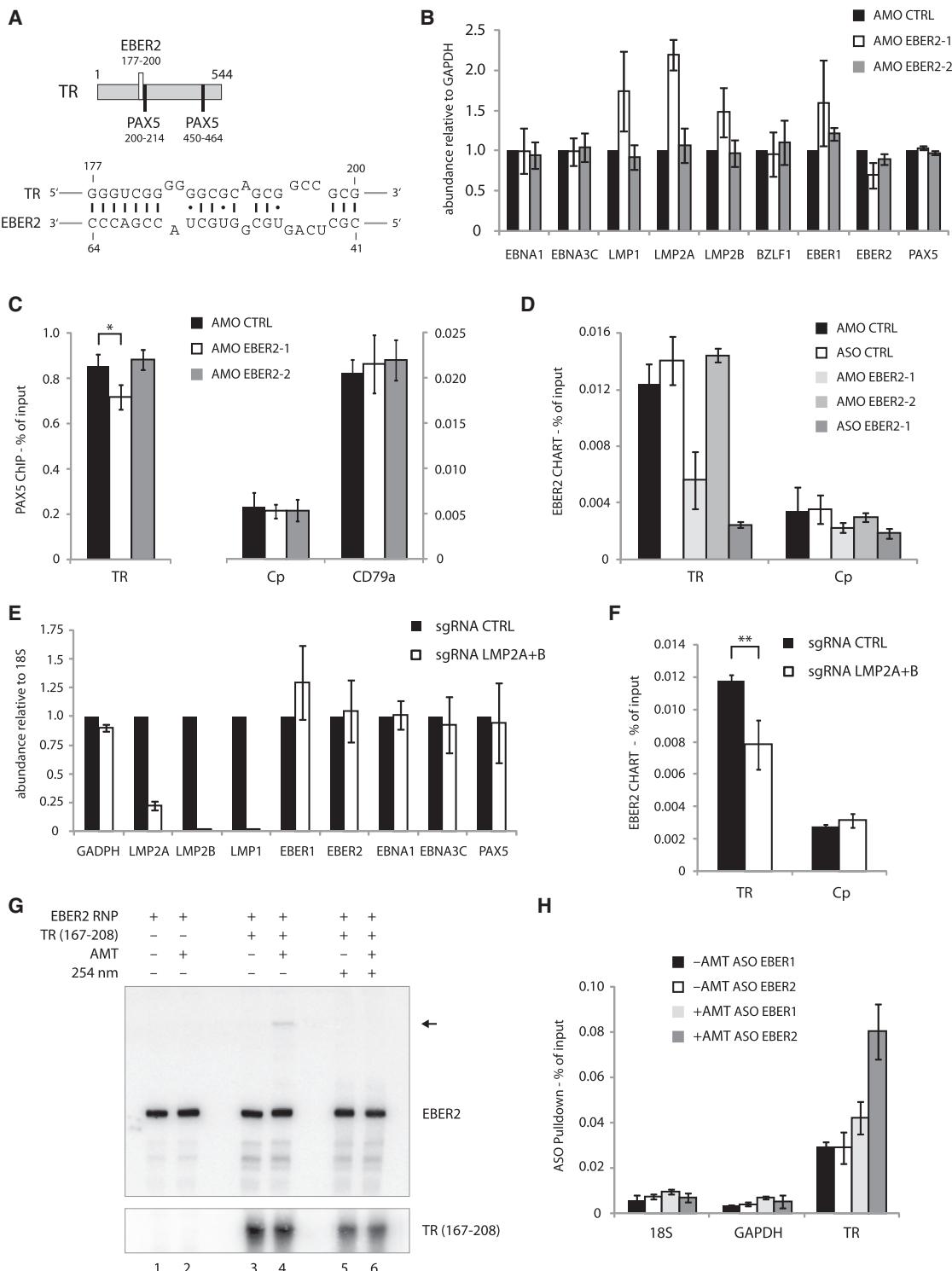
The TR sequence in EBV DNA contains two strong PAX5 consensus sequences (Arvey et al., 2012), suggesting that recruitment of the EBER2-PAX5 complex could be achieved through the DNA binding domain of PAX5. To test whether EBER2 is localized to the TRs by virtue of its interaction with PAX5, we knocked down PAX5 using lentivirally expressed short hairpin RNAs (shRNAs) as described previously (Arvey et al., 2012). Efficient depletion of PAX5 in BJAB-B1 cells was confirmed by western blot (Figure 3A); the knockdown efficiency (to  $\sim 40\%$  of the original level) was sufficient to elicit upregulation of LMP2A as measured by qRT-PCR (Figure 3B), arguing that the transcriptional control imposed by EBER2-PAX5 was compromised under these conditions. As expected upon depletion, PAX5 localization assessed by ChIP was significantly reduced at both the TRs and the CD79a promoter region (Figure 3C). However, surprisingly, EBER2 localization at the TRs remained unaffected as measured by CHART coupled to qPCR analysis (Figure 3D). We conclude that EBER2 is required for PAX5 localization (Figure 2G), but PAX5 is dispensable for EBER2 recruitment to the TRs.

Since EBER2 appears to be the key recruiting entity of the EBER2-PAX5 RNP, we considered the possibility of EBER2 recruitment via an RNA-RNA interaction. The fact that EBER2 binding to the TRs as assessed by CHART was not affected when RNase H digestion preceded ASO selection argues against an RNA-DNA interaction (Figures S3A and S3B). On

the other hand, a region within EBER2 that base pairs with an RNA transcribed from the TR region, such as the nascent transcripts of the two LMP2 isoforms, which contain the TRs in their first intron, might exist (Figure S4B). If this RNA-RNA-mediated recruitment model were correct, we should be able to (1) identify the complementarity, (2) show that transcription through the TRs is required for EBER2 recruitment, and (3) obtain evidence for a physical interaction between EBER2 and the nascent TR sequence-containing transcript *in vivo*.

Potential base pairing between EBER2 and the EBV type II TR RNA sequence was examined using the RNAup program to search for short, stable RNA-RNA interactions (Mückstein et al., 2006). A putative 18-bp hybrid was identified with a predicted free energy of  $\Delta G = -28.10$  kcal/mol. This hybrid comprises EBV TR nucleotides 177–200 and EBER2 nucleotides 41–64 (Figure 4A). Intriguingly, this sequence in EBER2 coincides with one of two oligonucleotide-accessible sites identified by RNase H digestion (Figure 1B). Notably, the sequence within the TR transcript (nucleotides 177–200) predicted to base pair with EBER2 lies adjacent to one of the PAX5 consensus sites (nucleotides 200–214; Figure 4A). A second putative 17-bp RNA-RNA hybrid of comparable predicted stability was identified that also overlaps with the same oligonucleotide-accessible region within EBER2 (site A versus site B of EBV type II; Figures S5B–S5D).

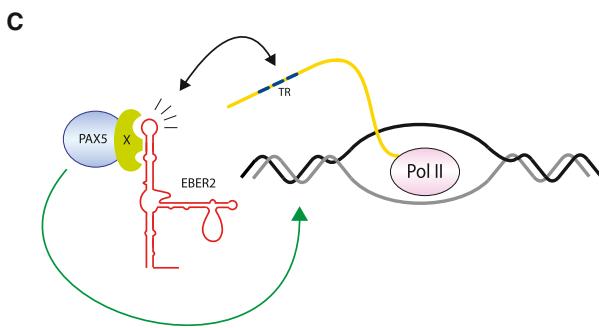
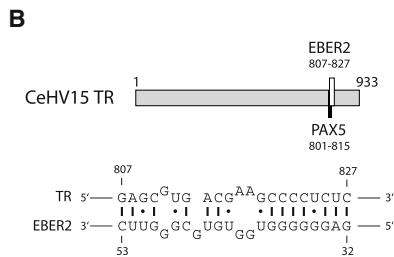
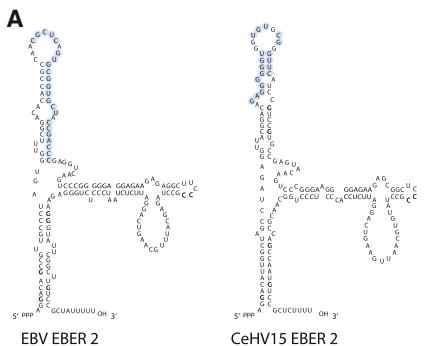
To show that the accessible region in EBER2 (nucleotides 41–64) is indeed necessary for its recruitment to TRs, we used an antisense morpholino oligonucleotide (AMO) that anneals to this region (AMO EBER2-1; Table S1; Figure 5A, left panel, nucleotides in blue) to block the putative EBER2-TR base-pairing interaction. We analyzed mRNA levels for several EBV-latent genes upon nucleofection of the 25-nt AMO EBER2-1 and observed that LMP2A, and to a lesser extent LMP1 and LMP2B, transcripts increased in level compared to a scrambled control AMO (AMO CTRL) and an AMO targeting EBER2 nucleotides 146–170 (AMO EBER2-2; Table S1), which are not predicted to form RNA-RNA interactions (Figures 4B and S3C). Moreover, AMO EBER2-1 treatment reduced PAX5 localization specifically at the TRs as measured by ChIP (Figure 4C). We were unable to use an AMO against the TR nucleotides 177–200 to block base



**Figure 4. EBER2 Is Recruited to the TRs through Base Pairing with Nascent RNA from the TR Locus**

(A) Predicted RNA-RNA interaction between EBER2 and a region within the TR (bottom). TR coordinates are also shown for the two PAX5 consensus sites (top). (B) RNA levels of several EBV genes were measured by qRT-PCR after 3 days of treatment with EBER2 AMO complementary to nucleotides 35–59 (AMO EBER2-1) or nucleotides 146–170 (AMO EBER2-2). (C) Quantification of the PAX5 ChIP at the TR, Cp, and CD79a loci after treatment with AMO EBER2-1 or EBER2-2. The control IgG ChIP data (data not shown) were comparable to those in Figure 2G. \* $p = 0.03$  (Student's t test,  $n = 3$ ). (D) ChIP analysis of EBER2 at TR and Cp loci after treatment with ASO EBER2-1 or EBER2-2. (E) RT-PCR analysis of nascent RNA levels of various genes under sgRNA CTRL or sgRNA LMP2A+B. (F) ChIP analysis of EBER2 at TR and Cp loci after treatment with sgRNA CTRL or sgRNA LMP2A+B. (G) RT-PCR analysis of EBER2 and TR (167–208) expression. (H) ASO pulldown assay of EBER2 and TR genes under four conditions: -AMT ASO EBER1, -AMT ASO EBER2, +AMT ASO EBER1, and +AMT ASO EBER2.

(legend continued on next page)



pairing of the nascent transcript with EBER2, as the nucleotide sequence of this AMO exhibits strong self-complementarity (data not shown). In summary, blocking the putative base-pairing region of EBER2 with an AMO results in the same phenotype as EBER2 depletion (Figures 2E and 2G). Importantly, CHART confirmed decreased EBER2 binding to the TRs in the presence of the AMO EBER2-1 that targets the predicted RNA-RNA interaction site (Figure 4D). These results argue that base pairing of EBER2 to the nascent LMP2A/B transcript could be instrumental in recruiting EBER2 and PAX5 to the TRs.

To demonstrate that transcription through the TRs is necessary for EBER2 recruitment, we interfered with LMP2 gene expression, which generates nascent transcripts containing TR sequences within the first intron (Figure 5C). We used CRISPR-mediated transcriptional interference by the catalytically inactive dCas9 protein to specifically silence both LMP2 isoforms (Gilbert et al., 2013). We generated stable dCas9-KRAB-expressing BJAB-B1 cells and identified potent single-guide RNAs (sgRNAs)

### Figure 5. EBER2-Guided Recruitment of PAX5 to the TRs Appears to Be Evolutionarily Conserved in a Related Gamma-Herpesvirus

(A) Secondary structure models of EBV EBER2 and the EBER2 homolog of rhesus LCV (CeHV15). Nucleotides predicted to base pair with TR RNA are indicated in blue.

(B) Relative position and coordinates within the TR, as well as the PAX5 consensus site, are shown (top). The predicted RNA-RNA interaction between CeHV15 EBER2 and the TR RNA is shown at the bottom.

(C) Model for complementary base-pairing-mediated recruitment of EBER2-PAX5 RNP to the TR region. See also Figure S3.

against LMP2A and LMP2B (Figures 4E and S4A–S4C). LMP1 expression was collaterally silenced by this approach, probably because its promoter is located close to the LMP2B transcription start site. Inhibiting the expression of nascent transcripts containing TR sequences by silencing both LMP2 isoforms resulted in decreased EBER2 binding to the TRs as determined by CHART (Figure 4F). Simultaneous usage of sgRNAs against both LMP2 isoforms was necessary to observe decreased

EBER2 binding, as inhibiting one isoform alone did not affect EBER2 localization to the TRs (Figure S4D; data not shown).

Finally, to confirm *in vivo* the predicted base-pairing interaction between EBER2 and TR sequence-containing RNA, we used the psoralen derivative aminomethyltrioxsalen (AMT). AMT preferentially crosslinks pyrimidine bases on opposite strands of double-stranded RNA after irradiation with 365 nm UV light; 254 nm UV light irradiation reverses the crosslinks (Cimino et al., 1985). First, we ascertained that AMT is able to crosslink the predicted interaction sites between EBER2 and nascent RNA containing the TR sequence. We *in vitro* transcribed a 42-nt RNA from the TR region (Table S1) that contains the sequence predicted to base pair with EBER2 and added it to an EBER2 RNP containing cell lysate. Only in the presence of AMT and the *in-vitro*-transcribed TR RNA did we observe a higher molecular weight band in an EBER2 Northern blot (Figure 4G, lane 4, arrow). This band disappeared when crosslinks were reversed by irradiating with 254 nm UV light (Figure 4G, lane 6), indicating that the predicted base pairs form

(D) EBER2-CHART followed by qPCR analysis after AMO EBER2-1 or KD ASO treatment.

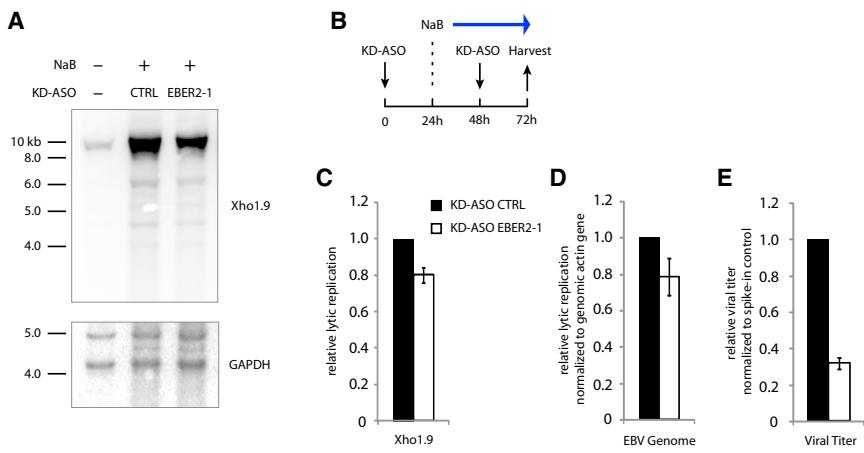
(E) qRT-PCR analysis after expressing sgRNAs targeting LMP2A and LMP2B (Figure S4B) in dCas9-KRAB-expressing BJAB-B1 cells.

(F) EBER2 CHART was conducted after CRISPR/dCas9-mediated transcriptional interference of LMP2 genes. \*\*p = 0.01 (Student's t test, n = 3).

(G) EBER2-RNP-containing cell lysate was incubated with an *in-vitro*-transcribed 42-nt RNA fragment from the TR region (nucleotides 167–208) predicted to base pair with EBER2 as shown in (A). AMT was added where indicated, and the reaction was exposed to long-wave UV light (365 nm). Short-wave UV-light irradiation (254 nm) was included as indicated to reverse crosslinks. RNA was isolated, and northern blotting was carried out on a denaturing urea-polyacrylamide gel, probing for EBER2 and the *in-vitro*-transcribed TR RNA. An arrow indicates EBER2 crosslinked to the TR RNA fragment.

(H) EBV-positive cells were treated with long-wave UV light in the presence or absence of AMT. EBER1 and EBER2 were selected using specific ASOs and, together with co-precipitated RNAs, were reverse transcribed for qRT-PCR analysis. The abundance of TR-sequence-containing RNA was measured in EBER1- and EBER2-selected samples. Primers detecting 18S rRNA and GAPDH mRNA were used as negative controls. All data represent the mean of three independent experiments ± SD.

See also Figures S3 and S4 and Table S1.



**Figure 6. EBER2 Depletion Results in Decreased Viral Lytic Replication**

(A) Viral lytic replication after KD-ASO-mediated EBER2 depletion was measured by Southern blot analysis using the Xho1.9 probe. The same blot was probed for the cellular GAPDH locus as a loading control.

(B) Experimental outline for viral lytic induction by sodium butyrate (NaB) in combination with EBER2 knockdown by KD ASO.

(C) Quantification of three independent experiments as shown in (A).

(D) Viral lytic replication was measured by qPCR analysis, amplifying the EBV dyad symmetry DNA region normalized to the cellular actin gene.

(E) Relative viral titer in supernatant was measured by qPCR normalized to spike-in control. All data represent the mean of three independent experiments  $\pm$  SD.

See also Figure S6 and Table S1.

in vitro. In vitro crosslinking was also observed with the second EBER2-TR interaction site (Figure S5E), suggesting that both RNA-RNA interactions could contribute to association in vivo. To show that base-pairing interaction does occur in vivo, we enriched EBER2, as well as EBER1 as a control, using ASO beads under denaturing conditions from a lysate of cells after in vivo psoralen crosslinking. EBERs, together with crosslinked RNAs, were eluted from the ASO beads with tetraethylammonium chloride-containing buffer to minimize background (Figure S5A). TR RNA was enriched by EBER2 compared to EBER1 ASO beads only after psoralen crosslinking as measured by qRT-PCR (Figure 4H). No enrichment was observed for 18S rRNA or GAPDH mRNA. Together, these results argue that EBER2 base pairs with TR sequence-containing nascent transcripts of the LMP2A and 2B genes in vivo.

#### EBER2's Interaction with Nascent RNA Is Evolutionarily Conserved

Of the many EBV-related primate LCVs (Lacoste et al., 2010), complete genome sequences are available only for EBV (type I and type II), for the rhesus (*Cercopithecine herpesvirus 15*, CeHV15), and for the marmoset LCV (*Callitrichine herpesvirus 3*) (Rivaiiller et al., 2002a, 2002b). The genome of the last has apparently lost its EBER2 gene, while CeHV15 retains an EBER2 homolog. Even though only moderate sequence conservation (65%) is exhibited, CeHV15 EBER2 can be modeled to fold into a structure that is almost identical to that of EBV EBER2 (Figure 5A). The CeHV15 TR, on the other hand, has no obvious sequence similarity to the EBV TR sequence, except for high overall GC content and a tandem repeat organization (Rivaiiller et al., 2002b); the repeat unit is considerably longer (933 bp) and contains only a single PAX5 consensus sequence (Figure 5B, top).

Using the RNAup program, we searched for an RNA-RNA interaction between CeHV15 EBER2 and its TR sequence. Only one stable hybrid was predicted with a free energy of  $\Delta G = -25.60$  kcal/mol (Figure 5B, bottom). Strikingly, the sequence within the CeHV15 EBER2 homolog predicted to base pair with the TR is in the same relative location as is that in EBV EBER2 (Figure 5A, nucleotides in blue). Furthermore,

despite great sequence divergence, the region within CeHV15 TR predicted to base pair with its EBER2 homolog overlaps the PAX5 consensus site, similar to its position in the EBV TR (compare Figure 5B and Figure 4A). Thus, the EBER2 guide function of PAX5 to the TRs appears to be evolutionarily conserved in the rhesus virus and possibly other yet unsequenced LCVs.

#### EBER2 Knockdown Affects Viral Lytic Replication

Given the localization of EBER2 at the TRs, which during the lytic cycle are the sites of viral genome linearization upon packaging into capsids (Zimmermann and Hammerschmidt, 1995), we asked whether viral lytic replication might be affected by EBER2 depletion. We treated the replication-permissive EBV-positive cell line HH514-16 with sodium butyrate (NaB) to induce lytic replication in combination with EBER2 KD ASO nucleofection (Figures 6B and S6A) (Ragoczy et al., 1998). Genomic DNA was isolated and subjected to Southern blot analysis using the Xho1.9 probe, which covers a unique sequence adjacent to the TR region, to quantify lytic replication (Figure S6B) (Raab-Traub and Flynn, 1986). After EBER2 depletion, lytic replication was decreased as analyzed by Southern blot and qPCR (Figures 6A, 6C, and 6D). The viral titer in the culture medium was reduced to an even greater extent (Figure 6E). Depletion of PAX5 consistently resulted in comparable decreased lytic replication (Figures S6C–S6E), supporting the notion that EBER2 and PAX5 act in concert. NaB promotes the expression of Zebra, an EBV transcription activator encoded by the BZLF1 gene that acts to induce lytic replication (Miller et al., 2007). EBER2 knockdown did not affect Zebra expression (Figure S6F), ruling out the possibility that the decreased lytic replication was due to a lower level of Zebra. Interestingly, the decrease in viral replication was restricted to lytic replication, since latent replication, measured in the absence of NaB induction (see Figure 2C), was unaffected upon EBER2 knockdown (Figure S6G).

#### DISCUSSION

Using CHART, we identified EBER2 as localized to the TRs of the latent EBV genome. We were unable to examine EBER1

localization by CHART because no region in EBER1 is accessible for hybridization with an ASO (Lee et al., 2012). A potential function of EBER2 at the TRs was suggested by its chromatin co-localization and interaction with the B cell master regulator PAX5. Even though two strong PAX5 consensus sequences are present in each TR unit, recruitment of the EBER2-PAX5 RNP does not appear to be achieved by direct PAX5 binding to DNA but is greatly facilitated by nucleotide sequence information contributed by EBER2 (Figure 5C). Perhaps because of the degenerate nature of the PAX5 consensus sequence (Figure S2D), the EBER2 RNP is required for efficient recruitment to the DNA target site (Figure 2). Alternatively, the EBER2 RNP may be required to clear adjacent chromatin of interfering factor(s) or to stabilize PAX5-DNA binding. We identified a base-pairing interaction between EBER2 and nascent transcripts from the TR regions that could provide such enhanced targeting specificity (Figure 4A). Indeed, the process of transcription through the TR regions per se might be integral to the recruitment mechanism by opening the chromatin conformation and thus facilitating accessibility of the PAX5-EBER2 RNP to its target site.

We identified PAX5 as a novel EBER2-interacting protein, prompted by the observation that this transcription factor and the viral noncoding RNA co-localize at the TRs (Figure 1). The interaction appears to be indirect, based on the negative results of EMSAs and UV crosslinking experiments (Figures S2C-S2G). We are currently attempting to further analyze the EBER2-PAX5 RNP, particularly focusing on the factor that bridges EBER2 and PAX5. Depletion of either EBER2 or PAX5 exhibits overlapping phenotypes, such as the upregulation of LMP genes (Figure 2E), suggesting both a functional and a physical interaction of the two factors. PAX5 depletion was reported to have a broader effect on EBV-latent gene expression (Arvey et al., 2012) than we observe here, which might be explained by the different time points and methods used. While Arvey et al. (2012) achieved PAX5 knockdown by lentiviral expression of shRNAs and gene expression changes were examined after 5 days, we chose an earlier time point (3 days) for analyzing gene expression upon EBER2 depletion by KD ASOs in an effort to exclude secondary and pleiotropic effects of long-term PAX5 knockdown.

Our attempts to recapitulate EBER2-PAX5-guided transcriptional silencing at EBV TRs in heterologous reporter systems were unsuccessful. When the entire LMP2A locus, including the TRs in its first intron, was cloned into a vector and expressed together with EBER2 and PAX5 in B lymphocytes, no effect on LMP2A expression was observed (Figure S3D). Similarly, incorporating TRs into the 3' end of a luciferase reporter gene was also unresponsive to the presence of EBER2 and PAX5 (Figure S3E). Apparently, a nascent transcript from the TRs alone is not sufficient for PAX5 recruitment through EBER2, and other factors, possibly DNA elements, are necessary to establish a silenced chromatin architecture.

PAX5 at the TRs appears to exert a function other than acting as a classical transcription activator or a repressor, as is normally found at promoters. The repeat organization of the TRs might be a crucial aspect of the PAX5 localization mechanism and perhaps also of function. Indeed, the related transcription factors, Pax3 and Pax9, have been reported to restrict RNA output from mouse satellite repeat sequences by binding and recruiting

histone methyl transferases to silence repetitive DNA (Bulut-Karslioglu et al., 2012). In another uninfected cellular context, Pax5 action has been reported to regulate the immunoglobulin heavy chain (*Igh*) locus during VDJ recombination in pro-B cells (Ebert et al., 2011). Parallels include (1) repeat regions bound by Pax5, the so-called Pax5-activated intergenic repeat elements (PAIRs), of which 14 are interspersed in the *Igh* locus, and (2) strikingly, the existence of a noncoding RNA expressed from the *Igh* locus. Expression of the RNA coincides with Pax5 binding to the PAIRs, whereas at later times in B cell development when the RNA ceases to be made, Pax5 localization is no longer detected. It is tempting to speculate that the *Igh* noncoding transcript might contribute to Pax5 recruitment similarly to the nascent transcript emanating from the TR regions of EBV. Furthermore, Pax5 induces chromatin condensation of the *Igh* locus (Fuxa et al., 2004). If an analogous chromatin contraction occurs at the TRs of the EBV genome, transcriptional upregulation of LMP genes following perturbation of the EBER2-PAX5-mediated control mechanism might be explained by a looser chromatin conformation that facilitates transcription through the region.

Consistent with the possibility that genome organization is regulated by the EBER2-PAX5 interaction, EBER2 depletion does not result in immediate transcriptional upregulation of EBV genes nearest to its binding site; changes become apparent only after 3 days of knockdown. This observation suggests that the genome organization at the TRs, once established, remains stable unless the correct organization cannot be resumed following genome replication and/or dilution of regulating factors by knockdown. Importantly, we demonstrate that EBER2 depletion affects viral lytic replication and propose that decreased lytic replication might be a consequence of improper genome organization that hinders efficient replication. Perhaps, latent replication is not affected because the TRs contribute differently to the replication of EBV episomes, compared to the production of linear packaged virion DNA. It is of course possible that there are additional consequences of EBER2 knockdown that we have not assessed in this study. One such possibility, given the fact that both EBER2 and lytic replication have been implicated in promoting oncogenicity (Katsumura et al., 2012; Ma et al., 2011), is an interplay between the function of EBER2 at the TRs and the consequences of lytic replication on tumor formation. An indication that EBER2 could have function(s) in addition to recruiting PAX5 to the TRs stems from the fact that more EBER2 molecules are present than complementary TR binding sites; in an infected cell, there are  $2.5 \times 10^5$  EBER2 molecules and up to 50 EBV episomes, each containing up to 20 TRs, although each TR probably harbors multiple nascent transcripts. On the other hand, the overabundance of EBER2 molecules compared to the number of TR binding sites during latency could be necessary to accommodate the massive increase in EBV genome copy number, and hence TR binding sites, occurring during viral lytic replication.

In recent years, more and more long noncoding RNAs (lncRNAs), arbitrarily defined as >200 nucleotides in size, have been shown to fulfill a diversity of cellular functions (Cech and Steitz, 2014). In addition to post-transcriptional regulation, a common theme is the interaction of lncRNAs with chromatin-modulating factors to control gene expression (Huarte et al.,

2010; Nagano et al., 2008; Tsai et al., 2010; Zhao et al., 2008). Thus, lncRNAs have been proposed in theory to act as targeting guides for effector proteins by base pairing with specific chromatin sites via RNA-RNA interactions, RNA-DNA interactions (Gilbert et al., 2013), or triplex formation (Schmitz et al., 2010). RNA-RNA interactions are the best-supported interaction mode experimentally for noncoding RNAs smaller than 200 nucleotides, as exemplified by small nucleolar RNAs acting in RNA modification, small nuclear RNAs in pre-mRNA splicing, or microRNAs in targeting mRNAs (Kim et al., 2009; Watkins and Bohnsack, 2012; Will and Lührmann, 2011). A guide function for targeting specific sites on chromatin thus far has been ascribed to tiny RNAs only, such as piwi-interacting RNA-mediated recruitment of PIWI in *Drosophila*, small-interfering-RNA (siRNA)-mediated centromeric silencing in yeast, and siRNA-directed DNA methylation in plants (Lejeune et al., 2010; Malone and Hannon, 2009; Matzke et al., 2007). All previously reported lncRNAs appear to fulfill an architectural scaffolding function, often with chromatin-regulating proteins. Here, for the first time, we provide evidence for a base-pairing interaction of a trans-acting moderately sized noncoding RNA, EBER2, which facilitates the recruitment of an associated transcription factor to chromatin target sites.

The ability of EBER2 to help recruit PAX5 to the TRs appears to be evolutionarily conserved not only in both type I and type II EBV (Figures S5B–S5E) but also in a related rhesus LCV. Marmoset LCV, a virus that infects a new world primate, has lost an EBER2 homolog but nonetheless retains a strong PAX5 consensus site within its TR (Rivailler et al., 2002a), suggesting that PAX5 binding occurs. This raises the question of whether PAX5 is recruited by an analogous mechanism involving a yet unidentified noncoding RNA, which has replaced EBER2, or whether a compensatory mechanism not involving an RNA-RNA interaction mediates recruitment of PAX5 in marmoset LCV. A better understanding of the precise EBER2-PAX5 RNP composition will be essential to distinguish between the two possibilities. Another open question is whether EBER2 is recruited to TRs in EBV-infected cells that do not express the B-cell-specific factor PAX5, such as nasopharyngeal carcinoma (NPC) cells, which are epithelial. As NPCs exhibit latency II (Rowe et al., 2009), characterized by robust expression of LMP1 and LMP2 genes, as well as EBER2, EBER2-PAX5-mediated transcription inhibition is unlikely to occur in these cells. Similarly, the questions of whether and how PAX5 is recruited to the TRs in EBV strains carrying a deletion of the EBER2 gene also remain to be answered.

## EXPERIMENTAL PROCEDURES

### CHART Assay

Endogenous RNase H cleavage assays were performed as described in Lee et al. (2012). CHART was carried out as described in Simon (2013), with minor modifications. CHART-seq data were deposited in the Sequence Read Archive under the accession number SRR1640963. For detailed protocols, see the Extended Experimental Procedures.

### EBER2 and PAX5 Knockdown

$2.5 \times 10^6$  BJAB-B1 cells were nucleofected with 10  $\mu$ l of 100 mM KD ASO/AMO stock solution in SF solution with program EN-150 using the Lonza

4D-Nucleofector System. HH514-16 cells were nucleofected using the Lonza 2b Device with solution V and program A-023. On the day following nucleofection, cells were separated from debris using Lymphocyte Separation Medium (Corning Cellgro) in accordance with the manufacturer's instructions. For KD ASO and AMO sequence information, see Table S1. KD ASOs were designed (Ideue et al., 2009) with DNA nucleotides flanked by 2'-O-methyl nucleotides and consisting of a phosphorothioate backbone (for increased stability) to induce cleavage by endogenous RNase H.

RNAi against PAX5 was performed by lentiviral shRNA expression using MISSION shRNA clones TRC0000016061 and TRC0000016062 (SIGMA) as described previously (Arvey et al., 2012; Cozma et al., 2007). Lentiviruses were produced as described previously (Lee et al., 2012). Cells were cultured under puromycin selection 1 day after infection and harvested 4 days post-infection, as a decrease in proliferation rate became apparent at this time point.

### Psoralen Crosslinking of RNAs

A 42-nt or 36-nt RNA within the TR region (nucleotides 167–208 and 74–109 of EBV type II, respectively) predicted to base pair with EBER2 was in vitro transcribed with T7 polymerase from an oligonucleotide template (see Table S1 for sequence). For psoralen crosslinking of EBER2 RNP to the in-vitro-transcribed TR fragment, nuclei were isolated by lysing  $10^7$  BJAB-B1 cells in 10 mM Tris (pH 8.0), 0.32 M sucrose, 3 mM CaCl, 0.1 mM EDTA, and 0.1% NP-40 and were then resuspended in 100  $\mu$ l of 10 mM HEPES (pH 7.4), 150 mM KCl, 5 mM MgCl<sub>2</sub>, 0.2 mM DTT, 10% glycerol, and 0.5% NP-40 to generate nuclear extract. The lysate was cleared by a 3-min centrifugation step at full speed in a table-top centrifuge. 200 ng of in-vitro-transcribed TR fragment was added to 10  $\mu$ l of nuclear extract in the presence of 40  $\mu$ g/ml aminomethyltrioxsalen (AMT), 0.2  $\mu$ g tRNA, and 40 U RNase inhibitor and incubated 30 min at room temperature (RT) before the reaction was irradiated for 30 min on ice covered with a 2-mm-thick glass plate from a distance of 2.5 cm with a handheld 365-nm UV lamp. Extracts were treated with 254-nm UV irradiation for 10 min on ice, where applicable. After crosslinking, RNA was isolated with TRIZOL and subjected to Northern blot analysis.

For in vivo crosslinking of intact cells,  $2 \times 10^7$  cells were resuspended in 1 ml of growth medium containing 50  $\mu$ g/ml AMT, incubated for 5 min at 37°C before chilling the cells on ice, and irradiated with UV light as described above. Cells were washed with PBS, and RNA was isolated with TRIZOL and DNase treated. 20  $\mu$ g total RNA was heated in 100  $\mu$ l TE buffer at 95°C for 3 min and chilled on ice before 50  $\mu$ l Denaturant buffer and 150  $\mu$ l 2× Hybridization buffer were added (see CHART protocol above). 25  $\mu$ l of biotinylated (EBER1 or EBER2) ASO-streptavidin Dynabeads were added and incubated overnight at RT. Beads were washed three times with CHART wash buffer and once with 2.4 M tetraethylammonium chloride (TEACl) at 25°C, and bound RNAs were eluted with 2.4 M TEACl for 5 min at 40°C, followed by phenol-chloroform extraction prior to Northern blot or qRT-PCR analyses.

### ACCESSION NUMBERS

The SRA accession number for the CHART-seq reported in this paper is SRR1640963.

### SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, six figures, and one table and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2015.01.015>.

### ACKNOWLEDGMENTS

We thank Drs. George Miller (Yale), Ayman El-Guindy (Yale), and Nancy Raab-Traub (University of North Carolina) for EBV reagents and discussions, Matthew Simon (Yale), Paul Lieberman (Wistar), Tom Cech (UC Boulder), Thomas Jenuwein (MPI Freiburg), Sandra Weller (UConn Health), and Joel Rzowsky (Yale) for advice, and Jessica Brown, Kazimierz Tycowski, and Mingyi Xie for helpful comments. This work was supported by a grant from the

National Cancer Institute (CA16038). J.A.S. is an investigator of the Howard Hughes Medical Institute.

Received: August 21, 2014

Revised: November 14, 2014

Accepted: January 6, 2015

Published: February 5, 2015

## REFERENCES

Arvey, A., Tempera, I., Tsai, K., Chen, H.S., Tikhmyanova, N., Klichinsky, M., Leslie, C., and Lieberman, P.M. (2012). An atlas of the Epstein-Barr virus transcriptome and epigenome reveals host-virus regulatory interactions. *Cell Host Microbe* 12, 233–245.

Brown, N.A., Liu, C.R., Wang, Y.F., and Garcia, C.R. (1988). B-cell lymphoproliferation and lymphomagenesis are associated with clonotypic intracellular terminal regions of the Epstein-Barr virus. *J. Virol.* 62, 962–969.

Bulut-Karslioglu, A., Perrera, V., Scaranaro, M., de la Rosa-Velazquez, I.A., van de Nobelen, S., Shukeir, N., Popow, J., Gerle, B., Opravil, S., Pagani, M., et al. (2012). A transcription factor-based mechanism for mouse heterochromatin formation. *Nat. Struct. Mol. Biol.* 19, 1023–1030.

Cech, T.R., and Steitz, J.A. (2014). The noncoding RNA revolution-trashing old rules to forge new ones. *Cell* 157, 77–94.

Chiu, S.H., Wu, M.C., Wu, C.C., Chen, Y.C., Lin, S.F., Hsu, J.T., Yang, C.S., Tsai, C.H., Takada, K., Chen, M.R., and Chen, J.Y. (2014). Epstein-Barr virus BALF3 has nuclease activity and mediates mature virion production during the lytic cycle. *J. Virol.* 88, 4962–4975.

Cimino, G.D., Gamper, H.B., Isaacs, S.T., and Hearst, J.E. (1985). Psoralens as photoactive probes of nucleic acid structure and function: organic chemistry, photochemistry, and biochemistry. *Annu. Rev. Biochem.* 54, 1151–1193.

Cozma, D., Yu, D., Hodawadekar, S., Azvolinsky, A., Grande, S., Tobias, J.W., Metzgar, M.H., Paterson, J., Erikson, J., Marafioti, T., et al. (2007). B cell activator PAX5 promotes lymphomagenesis through stimulation of B cell receptor signaling. *J. Clin. Invest.* 117, 2602–2610.

Ebert, A., McManus, S., Tagoh, H., Medvedovic, J., Salvagiotto, G., Novatchkova, M., Tamir, I., Sommer, A., Jaritz, M., and Busslinger, M. (2011). The distal V(H) gene cluster of the IgH locus contains distinct regulatory elements with Pax5 transcription factor-dependent activity in pro-B cells. *Immunity* 34, 175–187.

Feederle, R., Shannon-Lowe, C., Baldwin, G., and Delecluse, H.J. (2005). Defective infectious particles and rare packaged genomes produced by cells carrying terminal-repeat-negative Epstein-Barr virus. *J. Virol.* 79, 7641–7647.

Fok, V., Friend, K., and Steitz, J.A. (2006). Epstein-Barr virus noncoding RNAs are confined to the nucleus, whereas their partner, the human La protein, undergoes nucleocytoplasmic shuttling. *J. Cell Biol.* 173, 319–325.

Fuxa, M., Skok, J., Souabni, A., Salvagiotto, G., Roldan, E., and Busslinger, M. (2004). Pax5 induces V-to-DJ rearrangements and locus contraction of the immunoglobulin heavy-chain gene. *Genes Dev.* 18, 411–422.

Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Brar, G.A., Torres, S.E., Stern-Ginossar, N., Brandman, O., Whitehead, E.H., Doudna, J.A., et al. (2013). CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* 154, 442–451.

Gregorovic, G., Bosshard, R., Karstegl, C.E., White, R.E., Pattle, S., Chiang, A.K., Dittrich-Breiholz, O., Kracht, M., Russ, R., and Farrell, P.J. (2011). Cellular gene expression that correlates with EBER expression in Epstein-Barr virus-infected lymphoblastoid cell lines. *J. Virol.* 85, 3535–3545.

Greifenegger, N., Jäger, M., Kunz-Schughart, L.A., Wolf, H., and Schwarzmann, F. (1998). Epstein-Barr virus small RNA (EBER) genes: differential regulation during lytic viral replication. *J. Virol.* 72, 9323–9328.

Houmani, J.L., Davis, C.I., and Ruf, I.K. (2009). Growth-promoting properties of Epstein-Barr virus EBER-1 RNA correlate with ribosomal protein L22 binding. *J. Virol.* 83, 9844–9853.

Howe, J.G., and Shu, M.D. (1988). Isolation and characterization of the genes for two small RNAs of herpesvirus papio and their comparison with Epstein-Barr virus-encoded EBER RNAs. *J. Virol.* 62, 2790–2798.

Huarte, M., Guttmann, M., Feldser, D., Garber, M., Koziol, M.J., Kenzelmann-Broz, D., Khalil, A.M., Zuk, O., Amit, I., Rabani, M., et al. (2010). A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. *Cell* 142, 409–419.

Ideue, T., Hino, K., Kitao, S., Yokoi, T., and Hirose, T. (2009). Efficient oligonucleotide-mediated degradation of nuclear noncoding RNAs in mammalian cultured cells. *RNA* 15, 1578–1587.

Katsumura, K.R., Maruo, S., and Takada, K. (2012). EBV lytic infection enhances transformation of B-lymphocytes infected with EBV in the presence of T-lymphocytes. *J. Med. Virol.* 84, 504–510.

Kim, V.N., Han, J., and Siomi, M.C. (2009). Biogenesis of small RNAs in animals. *Nat. Rev. Mol. Cell Biol.* 10, 126–139.

Komano, J., Maruo, S., Kurozumi, K., Oda, T., and Takada, K. (1999). Oncogenic role of Epstein-Barr virus-encoded RNAs in Burkitt's lymphoma cell line Akata. *J. Virol.* 73, 9827–9831.

Lacoste, V., Lavergne, A., de Thoisy, B., Pouliquen, J.F., and Gessain, A. (2010). Genetic diversity and molecular evolution of human and non-human primate Gammaherpesvirinae. *Infect. Genet. Evol.* 10, 1–13.

Lee, N., Pimienta, G., and Steitz, J.A. (2012). AU1 hnRNP D is a novel protein partner of the EBER1 noncoding RNA of Epstein-Barr virus. *RNA* 18, 2073–2082.

Lejeune, E., Bayne, E.H., and Allshire, R.C. (2010). On the connection between RNAi and heterochromatin at centromeres. *Cold Spring Harb. Symp. Quant. Biol.* 75, 275–283.

Lerner, M.R., Andrews, N.C., Miller, G., and Steitz, J.A. (1981). Two small RNAs encoded by Epstein-Barr virus and complexed with protein are precipitated by antibodies from patients with systemic lupus erythematosus. *Proc. Natl. Acad. Sci. USA* 78, 805–809.

Lindahl, T., Adams, A., Bjursell, G., Bornkamm, G.W., Kaschka-Dierich, C., and Jahn, U. (1976). Covalently closed circular duplex DNA of Epstein-Barr virus in a human lymphoid cell line. *J. Mol. Biol.* 102, 511–530.

Ma, S.D., Hegde, S., Young, K.H., Sullivan, R., Rajesh, D., Zhou, Y., Jankowska-Gan, E., Burlingham, W.J., Sun, X., Gulley, M.L., et al. (2011). A new model of Epstein-Barr virus infection reveals an important role for early lytic viral protein expression in the development of lymphomas. *J. Virol.* 85, 165–177.

Malone, C.D., and Hannon, G.J. (2009). Molecular evolution of piRNA and transposon control pathways in *Drosophila*. *Cold Spring Harb. Symp. Quant. Biol.* 74, 225–234.

Matzke, M., Kanno, T., Huettel, B., Daxinger, L., and Matzke, A.J. (2007). Targets of RNA-directed DNA methylation. *Curr. Opin. Plant Biol.* 10, 512–519.

Medvedovic, J., Ebert, A., Tagoh, H., and Busslinger, M. (2011). Pax5: a master regulator of B cell development and leukemogenesis. *Adv. Immunol.* 111, 179–206.

Miller, C.L., Lee, J.H., Kieff, E., and Longnecker, R. (1994). An integral membrane protein (LMP2) blocks reactivation of Epstein-Barr virus from latency following surface immunoglobulin crosslinking. *Proc. Natl. Acad. Sci. USA* 91, 772–776.

Miller, G., El-Guindy, A., Countryman, J., Ye, J., and Gradoville, L. (2007). Lytic cycle switches of oncogenic human gammaherpesviruses. *Adv. Cancer Res.* 97, 81–109.

Moss, W.N., and Steitz, J.A. (2013). Genome-wide analyses of Epstein-Barr viruses reveal conserved RNA structures and a novel stable intronic sequence RNA. *BMC Genomics* 14, 543.

Mückstein, U., Tafer, H., Hackermüller, J., Bernhart, S.H., Stadler, P.F., and Hofacker, I.L. (2006). Thermodynamics of RNA-RNA binding. *Bioinformatics* 22, 1177–1182.

Nagano, T., Mitchell, J.A., Sanz, L.A., Paurer, F.M., Ferguson-Smith, A.C., Feil, R., and Fraser, P. (2008). The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. *Science* 322, 1717–1720.

Raab-Traub, N., and Flynn, K. (1986). The structure of the termini of the Epstein-Barr virus as a marker of clonal cellular proliferation. *Cell* 47, 883–889.

Ragoczy, T., Heston, L., and Miller, G. (1998). The Epstein-Barr virus Rta protein activates lytic cycle genes and can disrupt latency in B lymphocytes. *J. Virol.* 72, 7978–7984.

Repellin, C.E., Tsimbouri, P.M., Philbey, A.W., and Wilson, J.B. (2010). Lymphoid hyperplasia and lymphoma in transgenic mice expressing the small non-coding RNA, EBER1 of Epstein-Barr virus. *PLoS ONE* 5, e9092.

Revilla-I-Domingo, R., Bilic, I., Vilagos, B., Tagoh, H., Ebert, A., Tamir, I.M., Smeenk, L., Trupke, J., Sommer, A., Jaritz, M., and Busslinger, M. (2012). The B-cell identity factor Pax5 regulates distinct transcriptional programmes in early and late B lymphopoiesis. *EMBO J.* 31, 3130–3146.

Rivaillet, P., Cho, Y.G., and Wang, F. (2002a). Complete genomic sequence of an Epstein-Barr virus-related herpesvirus naturally infecting a new world primate: a defining point in the evolution of oncogenic lymphocryptoviruses. *J. Virol.* 76, 12055–12068.

Rivaillet, P., Jiang, H., Cho, Y.G., Quink, C., and Wang, F. (2002b). Complete nucleotide sequence of the rhesus lymphocryptovirus: genetic validation for an Epstein-Barr virus animal model. *J. Virol.* 76, 421–426.

Rovedo, M., and Longnecker, R. (2007). Epstein-Barr virus latent membrane protein 2B (LMP2B) modulates LMP2A activity. *J. Virol.* 81, 84–94.

Rowe, M., Kelly, G.L., Bell, A.I., and Rickinson, A.B. (2009). Burkitt's lymphoma: the Rosetta Stone deciphering Epstein-Barr virus biology. *Semin. Cancer Biol.* 19, 377–388.

Schmitz, K.M., Mayer, C., Postepska, A., and Grummt, I. (2010). Interaction of noncoding RNA with the rDNA promoter mediates recruitment of DNMT3b and silencing of rRNA genes. *Genes Dev.* 24, 2264–2269.

Simon, M.D. (2013). Capture hybridization analysis of RNA targets (CHART). *Curr. Protoc. Mol. Biol.* Chapter 21, Unit 21.25.

Simon, M.D., Wang, C.I., Kharchenko, P.V., West, J.A., Chapman, B.A., Alekseyenko, A.A., Borowsky, M.L., Kuroda, M.I., and Kingston, R.E. (2011). The genomic binding sites of a noncoding RNA. *Proc. Natl. Acad. Sci. USA* 108, 20497–20502.

Sugden, B., Phelps, M., and Domoradzki, J. (1979). Epstein-Barr virus DNA is amplified in transformed lymphocytes. *J. Virol.* 31, 590–595.

Swaminathan, S., Tomkinson, B., and Kieff, E. (1991). Recombinant Epstein-Barr virus with small RNA (EBER) genes deleted transforms lymphocytes and replicates in vitro. *Proc. Natl. Acad. Sci. USA* 88, 1546–1550.

Toczkowski, D.P., Matera, A.G., Ward, D.C., and Steitz, J.A. (1994). The Epstein-Barr virus (EBV) small RNA EBER1 binds and relocates ribosomal protein L22 in EBV-infected human B lymphocytes. *Proc. Natl. Acad. Sci. USA* 91, 3463–3467.

Tsai, M.C., Manor, O., Wan, Y., Mosammaparast, N., Wang, J.K., Lan, F., Shi, Y., Segal, E., and Chang, H.Y. (2010). Long noncoding RNA as modular scaffold of histone modification complexes. *Science* 329, 689–693.

Watkins, N.J., and Bohnsack, M.T. (2012). The box C/D and H/ACA snoRNPs: key players in the modification, processing and the dynamic folding of ribosomal RNA. *Wiley Interdiscip. Rev. RNA* 3, 397–414.

Will, C.L., and Lührmann, R. (2011). Spliceosome structure and function. *Cold Spring Harb. Perspect. Biol.* 3, 3.

Yajima, M., Kanda, T., and Takada, K. (2005). Critical role of Epstein-Barr Virus (EBV)-encoded RNA in efficient EBV-induced B-lymphocyte growth transformation. *J. Virol.* 79, 4298–4307.

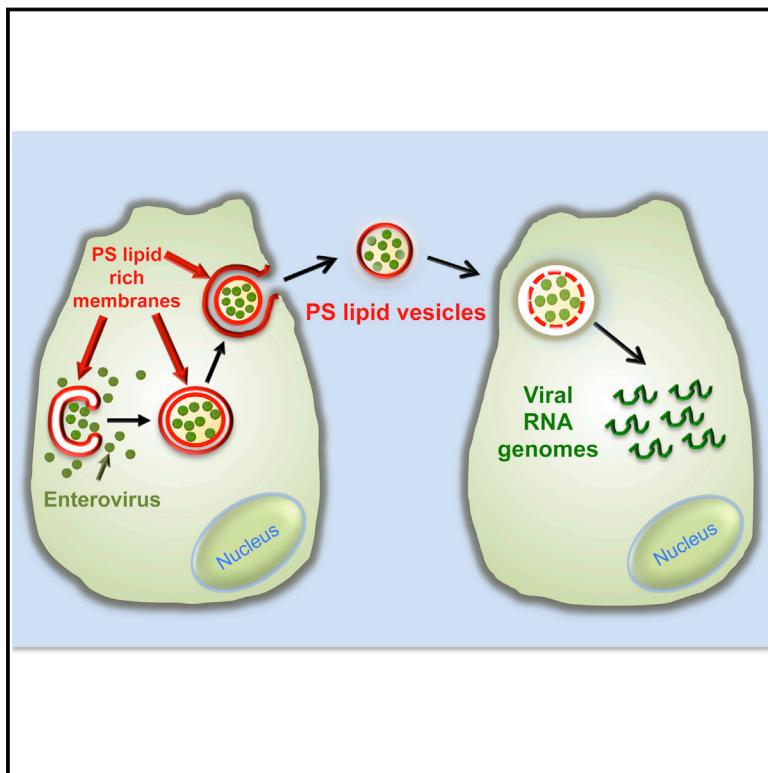
Young, L.S., and Rickinson, A.B. (2004). Epstein-Barr virus: 40 years on. *Nat. Rev. Cancer* 4, 757–768.

Zhao, J., Sun, B.K., Erwin, J.A., Song, J.J., and Lee, J.T. (2008). Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science* 322, 750–756.

Zimmermann, J., and Hammerschmidt, W. (1995). Structure and role of the terminal repeats of Epstein-Barr virus in processing and packaging of virion DNA. *J. Virol.* 69, 3147–3155.

# Phosphatidylserine Vesicles Enable Efficient En Bloc Transmission of Enteroviruses

## Graphical Abstract



## Authors

Ying-Han Chen, WenLi Du, ...,  
Grégoire Altan-Bonnet,  
Nihal Altan-Bonnet

## Correspondence

nihal.altan-bonnet@nih.gov

## In Brief

Clusters of enteroviruses are packaged in phosphatidylserine (PS)-enriched vesicles, thereby enhancing the infection efficiency of the viruses and enabling collective transmission of multiple viral genomes from cell-to-cell.

## Highlights

- Clusters of viruses are packaged and released non-lytically in PS lipid vesicles
- PS lipids are co-factors in mediating subsequent infectivity and transmission
- PS vesicles provide greater infection efficiency for viruses
- PS vesicles enable viral genome clusters to be transmitted en bloc cell-to-cell



# Phosphatidylserine Vesicles Enable Efficient En Bloc Transmission of Enteroviruses

Ying-Han Chen,<sup>1,2</sup> WenLi Du,<sup>1</sup> Marne C. Hagemeijer,<sup>1</sup> Peter M. Takvorian,<sup>2</sup> Cyrilla Pau,<sup>2</sup> Ann Cali,<sup>2</sup> Christine A. Brantner,<sup>3</sup> Erin S. Stempinski,<sup>3</sup> Patricia S. Connelly,<sup>3</sup> Hsin-Chieh Ma,<sup>4</sup> Ping Jiang,<sup>4</sup> Eckard Wimmer,<sup>4</sup> Grégoire Altan-Bonnet,<sup>5</sup> and Nihal Altan-Bonnet<sup>1,\*</sup>

<sup>1</sup>Laboratory of Host-Pathogen Dynamics, National Heart Lung and Blood Institute, NIH, Bethesda, MD 20892, USA

<sup>2</sup>Federated Department of Biological Sciences, Rutgers University, Newark, NJ 07102, USA

<sup>3</sup>Electron Microscopy Core Facility, National Heart Lung and Blood Institute, NIH, Bethesda, MD 20892, USA

<sup>4</sup>Department of Molecular Genetics and Microbiology, Stony Brook University, Stony Brook, NY 11794, USA

<sup>5</sup>Program in Computational Biology and Immunology, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA

\*Correspondence: nihal.altan-bonnet@nih.gov

<http://dx.doi.org/10.1016/j.cell.2015.01.032>

## SUMMARY

A central paradigm within virology is that each viral particle largely behaves as an independent infectious unit. Here, we demonstrate that clusters of enteroviral particles are packaged within phosphatidylserine (PS) lipid-enriched vesicles that are non-lytically released from cells and provide greater infection efficiency than free single viral particles. We show that vesicular PS lipids are co-factors to the relevant enterovirus receptors in mediating subsequent infectivity and transmission, in particular to primary human macrophages. We demonstrate that clustered packaging of viral particles within vesicles enables multiple viral RNA genomes to be collectively transferred into single cells. This study reveals a novel mode of viral transmission, where enteroviral genomes are transmitted from cell-to-cell en bloc in membrane-bound PS vesicles instead of as single independent genomes. This has implications for facilitating genetic cooperativity among viral quasispecies as well as enhancing viral replication.

## INTRODUCTION

Enteroviruses are a large genus of single positive-stranded RNA viruses whose members including poliovirus (PV), Coxsackievirus, rhinovirus, and enterovirus 68 are the causative agents of a number of important and widespread human diseases including poliomyelitis, myocarditis, hand foot and mouth disease, the common cold, and more recently, a severe respiratory disease with paralytic symptoms. In addition to >70 enteroviral serotypes identified in humans, enteroviral quasispecies are common largely as a result of inherent error making and lack of proofreading mechanisms of viral RNA-dependent RNA polymerases (RdRp).

Enteroviral RNA genomes serve as templates for both translation and replication, and these processes take place on host

intracellular membranes (den Boon and Ahlquist, 2010; Hsu et al., 2010). After enteroviruses have bound their specific host receptors either at the cell surface or within endocytic vesicles (Brandenburg et al., 2007), the capsid undergoes a conformational change that allows the viral RNA to be transferred across the plasma membrane and/or endosomal membrane into the cytoplasm through a yet completely defined mechanism (Strauss et al., 2013). In the cytoplasm, the enteroviral RNA is first translated into non-structural proteins and structural proteins, where the former makes up the RNA genome replication machinery and the latter the capsid. The viral RNA replication machinery are then assembled on the cytoplasmic membrane leaflet of ER-derived membranes that are subsequently modified by viral and host proteins to have a specific lipid blueprint of enrichment in phosphatidylinositol-4-phosphate (PI4P) and cholesterol lipids. These lipids regulate the membrane association, assembly, and activity of the viral replication protein complex, including the RdRp, and thus facilitate viral RNA synthesis (Hsu et al., 2010; Illytska et al., 2013; Nchoutmboue et al., 2013).

Once the enteroviral RNA is synthesized, little is known about where in the host cell it is packaged in capsids and how these capsids are released from cells. While enteroviruses have historically been considered non-enveloped (i.e., lacking a host-derived membrane bilayer around their capsids) and thus rely on cell lysis to exit, recent reports of extracellular Coxsackievirus B3 (CVB3) being present in vesicles (Robinson et al., 2014) and PV being able to spread non-lytically among host cells (Bird et al., 2014) have raised important questions regarding the extracellular nature of enteroviral particles and the significance of non-lytic exit in the viral life cycle. Moreover hepatitis A, hepatitis E and blue tongue viral particles, all long considered non-enveloped, have been observed surrounded by membranes (Feng et al., 2013; Takahashi et al., 2008; Owens et al., 2004).

A central paradigm in virology is that viruses behave as independent infectious units (Flint et al., 2009). While there are exceptions to this, such as vaccinia virus particles preventing superinfection by inducing the host cell to repel other virions (Doceul et al., 2010), it is largely accepted that the fate of individual viral genomes are not dependent on one another during exit from one cell and entry into another (Brandenburg and Zhuang

2007). Here, we investigate the assembly, exit, and subsequent infection processes of enteroviral particles using a combination of imaging techniques including confocal microscopy, super-resolution light microscopy, correlative light, and electron microscopy along with single molecule RNA fluorescence in situ hybridization (FISH), proteomic, and biochemical approaches. We show that multiple infectious enteroviral particles are clustered within individual phosphatidylserine (PS) lipid-enriched vesicles and non-lytically secreted out of cells. These viral particles in vesicles are more efficient in establishing infection than free viral particles. We demonstrate that vesicles encapsulate and traffic large numbers of mature infectious viral particles between cells and consequently enable the transfer of multiple viral RNA genomes collectively into new host cells by a mechanism that is dependent on both the virus-specific receptor of the recipient host cell as well as the vesicular PS lipids.

## RESULTS

### Assembled Poliovirus Capsids Are Localized to Viral RNA Replication Organelles

We first investigated the intracellular spatio-temporal dynamics of newly synthesized PV particles. The generation of PV particles, as well as enterovirus assembly in general, comprises a multistep process where capsid subunits (VP0, VP1, VP3) form pentamers, which polymerize into capsids (Liu et al., 2010). Once RNA is packaged, the VP0 subunits get cleaved into VP2 and VP4 to generate mature infectious virions (Liu et al., 2010). From screening a large number of neutralizing antibodies, we identified the A12 antibody, that binds deep inside the canyon bridging both rims of two adjacent pentamers and thus recognizing assembled capsids (Chen et al., 2011, 2013), to visualize PV capsids within infected cells. We fixed cells at various intervals after PV infection and co-immunolabeled with A12, anti-viral VP1 to detect individual VP1 subunits, and anti-viral 3AB antibodies, the latter to detect viral 3AB to localize replication organelles where viral RNA is synthesized (Hsu et al., 2010). Newly assembled capsids were clearly detectable from 3–4 hr post-infection (p.i.) and onward, and they were localized to the replication organelles (Figure 1A). Note that in contrast to A12 labeling, VP1 was localized to both the replication organelles and dispersed across the cytoplasm, consistent with its cytosolic nature. At 6–7 hr p.i., capsids were dispersed from the replication sites to the cytoplasm and sequestered in puncta (Figure 1A, 6–7 hr p.i. inset). At this time, there is a known cessation in viral RNA synthesis (Ehrenfeld et al., 1970), and the timing of capsid release from replication organelles was coupled to viral RNA synthesis since prematurely inhibiting viral RNA synthesis with the inhibitor Guanidine HCl (Barton and Flanagan, 1997) triggered capsids to rapidly disperse into the cytoplasm (Figure S1A).

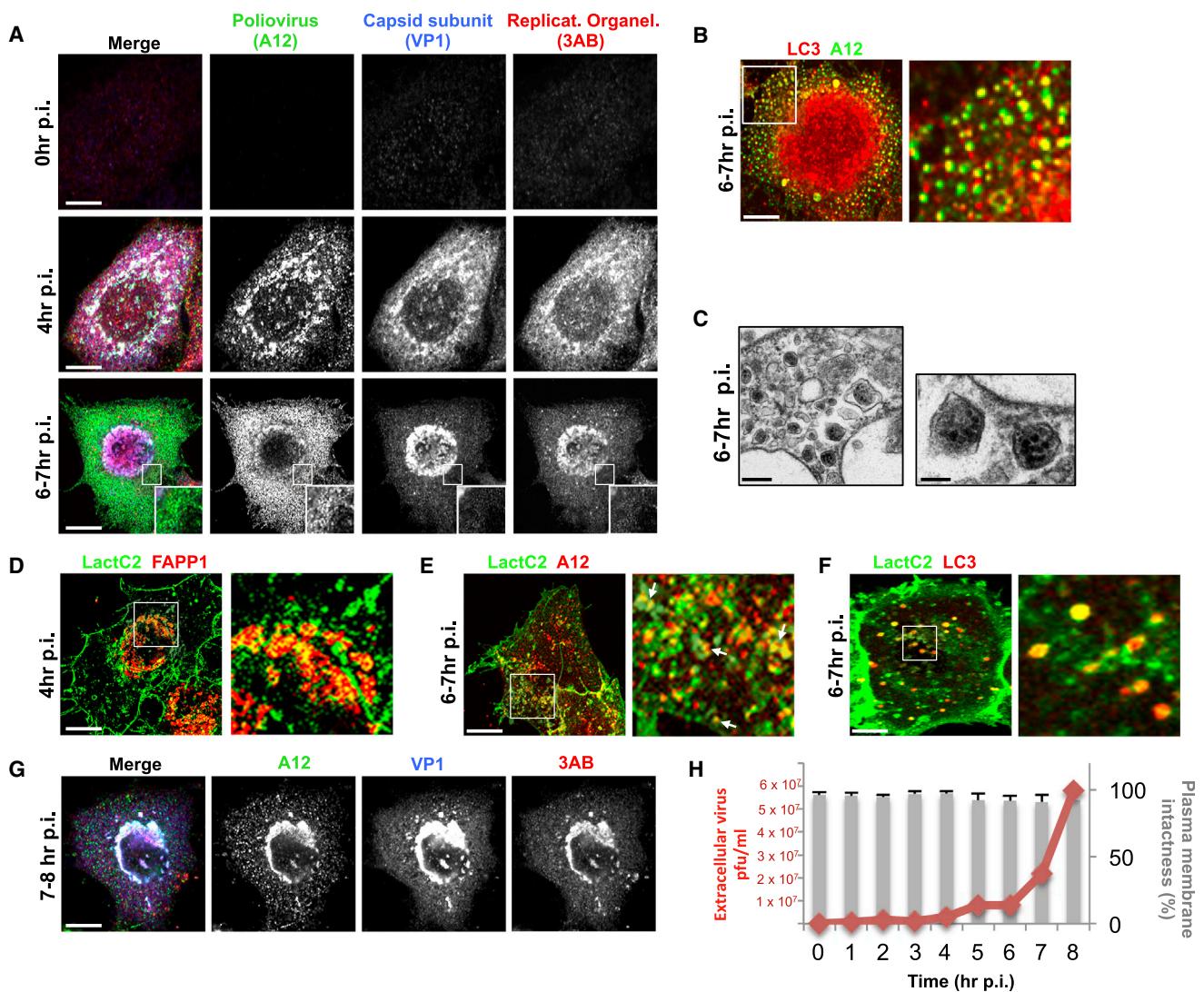
### Poliovirus Capsids Are Captured by Phosphatidylserine Lipid-Enriched Autophagosome-like Organelles and Released Non-Lytically from Cells

Between 6 and 7 hr p.i., we found that >85% of capsids ( $n = 10$  cells) were on punctate cytoplasmic structures that colocalized with the autophagosomal membrane protein, LC3-II (Figure 1B).

By transmission electron microscopy, we observed numerous double-membraned autophagosome-like organelles containing capsids (Figure 1C). Previous reports (Taylor et al., 2009; Kirkegaard and Jackson, 2005; Jackson et al., 2005) had found that perturbation of the host autophagy pathway led to a decrease in PV release from infected cells. Consistent with that, either disrupting autophagy by small interfering RNA (siRNA) depletion of autophagy machinery LC3 or beclin 1, or acutely stimulating autophagy by treating cells with tat-beclin 1 peptides, blocked or enhanced PV release by ~10-fold, respectively (Figures S1B and S1C) while replication was unaffected (Figure S1D). However, these capsid containing autophagosome-like organelles did not follow the conventional autophagy pathway and fuse their contents with lysosomes as inhibiting lysosomal enzymes did not further increase LC3-II levels beyond the 4-fold increase observed in PV-infected cells (Figures S1E and S1F) and none of the A12/LC3-II co-labeled structures contained lysosomal enzymes at any point during infection (Figures S1F and S1G). Notably, the SNARE protein syntaxin 17, normally localized to autophagosomes and required for fusion with lysosomes (Itakura et al., 2012), did not localize to these A12/LC3-II co-labeled structures (Figure S1I).

However, we found that the membranes of both the replication organelles and the autophagosome-like organelles contained negatively charged phosphatidylserine (PS) lipids. PS lipids in uninfected cells are primarily located at the cytoplasmic leaflets of the plasma and endosomal membranes as well as at the luminal leaflet of the ER membrane (Leventis and Grinstein, 2010; Kay et al., 2012). Cells were co-transfected with GFP-LactC2 and FAPP1mRFP, cytosolic live-cell reporters for PS (Kay et al., 2012), and PI4P lipids, respectively, the latter to report on the location of the ER-derived replication organelle membranes (Hsu et al., 2010). At 4 hr p.i., cells were imaged live by structured illumination microscopy. Highly localized PS-rich membrane domains exposed to the cytoplasm were found distributed across the replication organelles (Figure 1D, inset). Later, between 6 and 7 hr p.i., cells were fixed and immunolabeled with anti-GFP and A12 antibodies and imaged by confocal microscopy. Numerous A12 positive structures were found also co-labeled with GFP-LactC2 indicating the presence of PS lipids on their membrane leaflets exposed to the cytoplasm (Figure 1E, arrows). Consistent with this, in PV-infected live cells co-expressing LC3-mRFP and GFP-LactC2, >90% (organelles measured across ten cells) of the LC3-mRFP-labeled autophagosome-like organelles were co-labeled with GFP-LactC2 (Figure 1F, inset).

We then investigated the fate of these capsid-containing organelles during the rest of the infection time period. Between 7 hr and 8 hr p.i., there was a  $70\% \pm 10\%$  ( $n = 15$ ) decrease in the number of capsids within the cytoplasm (Figure 1G). Using the cell impermeable Trypan blue dye, we found that the plasma membrane remained intact during this time while there was an ~6-fold increase in extracellular viral titers (Figure 1H). This lack of plasma membrane permeability during PV infection had also been previously observed (Taylor et al., 2009; Bird et al., 2014). Although by 12 hr p.i. the cells eventually lysed (not shown), this data indicated that the majority of PV particles were released prior to cell lysis.



**Figure 1. Poliovirus Capsids Are Captured by Phosphatidylserine Lipid-Enriched Autophagosome-like Organelles and Released Non-Lytically from Cells**

(A) Capsids undergo dynamic spatial transitions during infection. HeLa cells infected with PV and immunolabeled with A12, anti-VP1, and anti-3AB antibodies. Scale bars represent 5 μm.

(B) Capsids (A12) colocalized with autophagosome marker LC3-II. PV-infected HeLa cells were immunolabeled with A12 and anti-LC3-II antibodies. Scale bar represents 5 μm.

(C) Electron micrographs of PV-infected cells show PV capsids in double-membrane autophagosome-like organelles. Scale bars represent 5 μm and 200 nm (inset).

(D) PV-infected cells at 4 hr p.i. expressing GFP-LactC2 and FAPP1mRFP imaged by structured illumination microscopy. Region of interest is magnified in right panel. Scale bar represents 5 μm.

(E) PV-infected cells at 6-7 hr p.i. expressing GFP-LactC2 were immunolabeled with anti-GFP and A12 antibodies. Region of interest is magnified in right panel. Arrows indicate A12 positive autophagosome-like organelles co-labeled with GFP-LactC2. Scale bar represents 5 μm.

(F) Cells co-expressing GFP-LactC2 and LC3-RFP were infected with PV and imaged by confocal microscopy at 7 hr p.i. Region of interest is magnified in right panel. Scale bar represents 5 μm.

(G) Capsid distribution between 7 and 8 hr p.i. PV-infected cells were immunolabeled with A12, anti-VP1, and anti-3AB antibodies. Scale bar represents 5 μm.

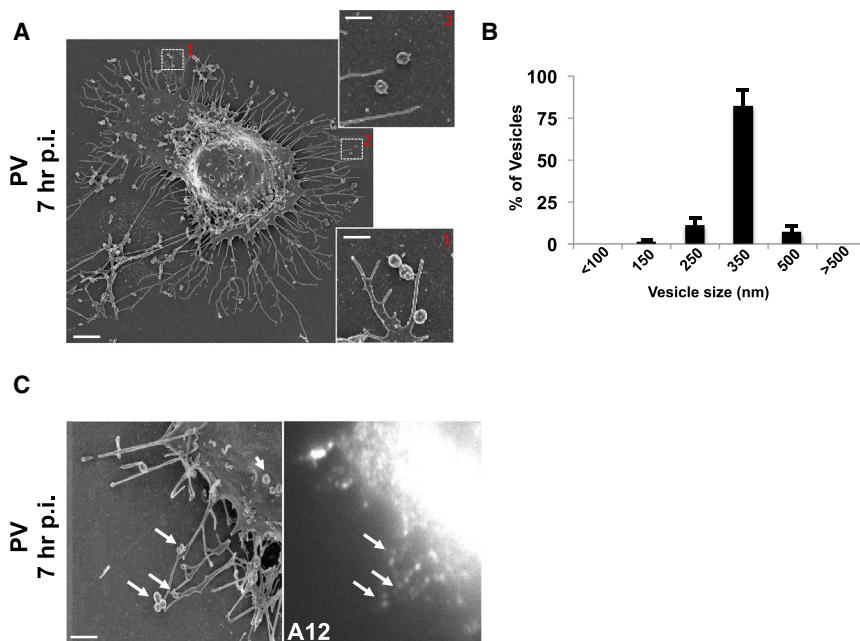
(H) Plasma membrane integrity remains intact when PV exits cells. Trypan Blue diffusion across the plasma membrane was measured concurrently with measurements of extracellular PV titer, the latter plotted in plaque-forming units/ml (pfu/ml).

See also Figure S1.

### Extracellular PV Particles Are Found in Uniformly Large Vesicles

Scanning electron microscopy (SEM) of cells at 7 hr p.i. cells revealed numerous vesicular structures of similar size docked at

the extracellular side of the intact plasma membrane (Figure 2A and inset). While these SEM images may not reflect the true shape and size of these vesicles within live cells, they do point to a striking uniformity in size. Measurements of the cross-section



**Figure 2. Extracellular PV Particles Are Found in Large Uniform-Sized Vesicles**

(A) Scanning electron microscopy of a PV-infected cell at 7 hr p.i. Scale bar represents 3  $\mu$ m. Inset shows higher magnification of uniform size vesicles docked on the extracellular side of the plasma membrane. Scale bar represents 1  $\mu$ m.

(B) Extracellular vesicle size distribution in PV-infected cells. Cross section diameter of a 100 randomly selected extracellular vesicles from four different cells, were measured from scanning electron micrographs and plotted. Data represented as mean  $\pm$  SEM.

(C) Correlative fluorescence and scanning electron microscopy (SEM). PV-infected cell was immunolabeled with A12 at 7 hr p.i., epifluorescence image was obtained (right) and then sample was processed for SEM (left). Arrows point to A12-labeled extracellular vesicles. Scale bar represents 1  $\mu$ m.

diameter from 100 randomly selected extracellular vesicles from four different cells, showed that ~90% of the vesicles were between 250 nm and 350 nm in diameter (Figure 2B). Correlative fluorescence imaging in conjunction with SEM confirmed that these vesicles contained A12-labeled capsids (Figure 2C).

#### Mature Enteroviral Particles Are Released in PS-Enriched Vesicles

We next investigated whether the extracellular vesicles containing PV retained the PS lipids that had been components of the autophagosome-like organelles (Figures 1D and 1F). Incubation of PV-infected cells with Alexa 568 coupled Annexin V, a non-cell permeable fluorescent reporter protein for PS lipids (Koopman et al., 1994), revealed numerous fluorescent puncta dotting the surface of the cell at 7 hr p.i. (Figure 3A). Similar results were also observed with CVB3 (at 7 hr p.i.) and human rhinovirus infections (at 12 hr p.i.) (Figure 3A). Note that this pattern of Annexin V labeling of enterovirus-infected cells was different from both apoptotic- and mock-infected cells: in the former, the entire plasma membrane was labeled with Annexin V as a result of PS lipids being on the extracellular membrane leaflet of the cell, a hallmark of apoptosis (Figure 3B, apoptosis) while in the latter there was no labeling since PS lipids were on the cytosolic leaflet of the plasma membrane (Figure 3B, mock).

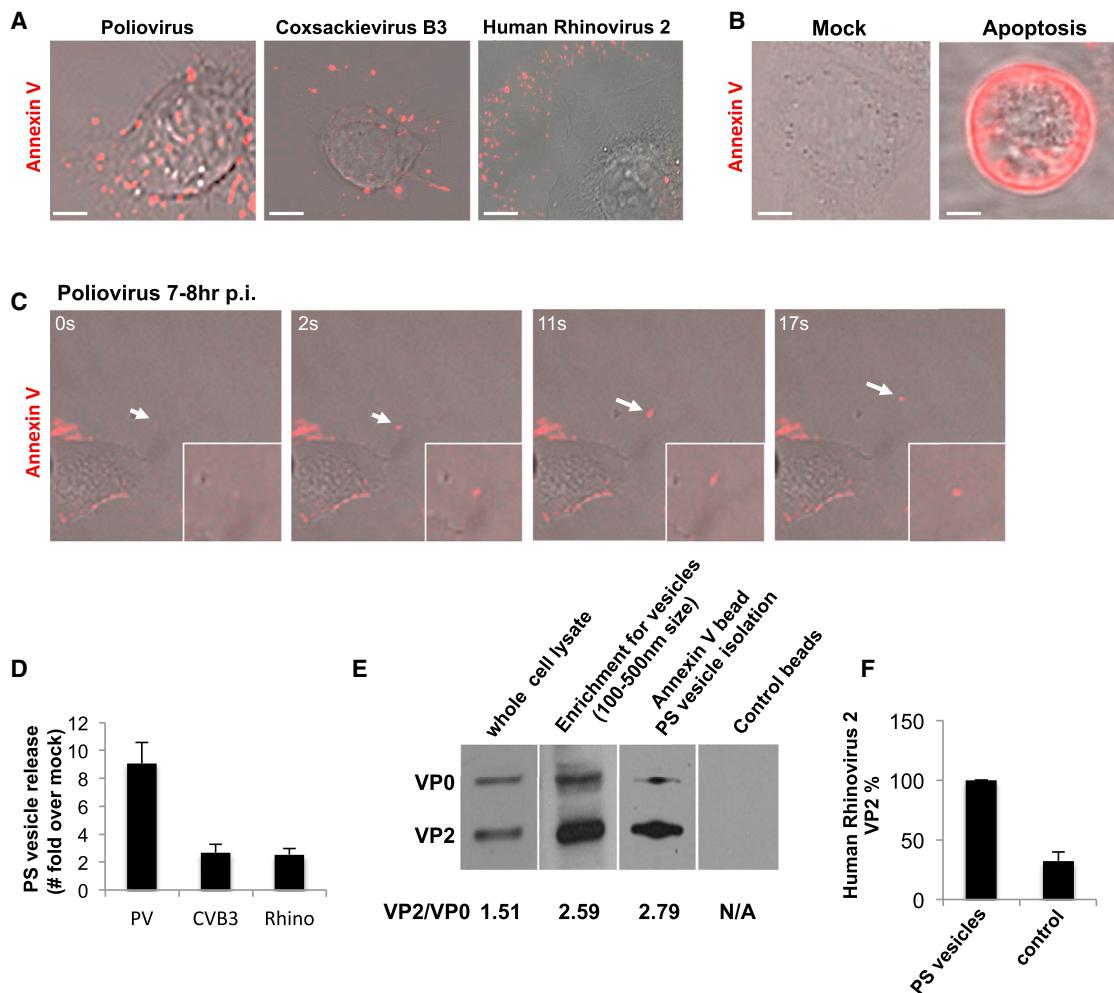
We performed time-lapse confocal/differential interference contrast (DIC) imaging on PV-infected cells in the presence of Alexa 568-Annexin V to determine if these vesicles were being released. We observed Annexin V-labeled vesicles emerging from the cell surface at discrete domains and being rapidly released into the extracellular medium (Figure 3C; arrow; Movie S1). We then quantified the amount of PS vesicles released during PV, CVB3, or rhinovirus infection relative to mock-infected cells (for each respective virus). We collected the extracellular medium, removed large cell debris, and enriched for vesicles

of size range 100–500 nm using differential centrifugation. This size range was chosen based on our previous light and electron microscopy data (Figure 2). The enrichment for vesicles of this size range was confirmed by transmission electron microscopy (Figure S2A). The vesicles were incubated with Alexa 568-Annexin V, and following the wash to remove any unbound Annexin V, placed in a spectrofluorometer. Fluorescence measurements revealed a net ~9-fold, ~3-fold, and ~3-fold increase in amounts of PS vesicles collected from the supernatants of PV-, CVB3-, and rhinovirus-infected cells, respectively, compared to mock-infected cells (Figure 3D).

We subsequently enriched for PS vesicles from the collected vesicle fraction using magnetic separation with Annexin V-conjugated magnetic microbeads. In parallel, vesicles were also incubated with magnetic microbeads lacking Annexin V to control for nonspecific binding. Post-magnetic separation, the samples were processed for SDS-PAGE/western analysis. We found that the extracellular vesicles both pre- and post-Annexin V isolation, had a VP2/VP0 ratio ~2-fold greater than the whole cell lysate (Figure 3E). This indicated that these extracellular released PS vesicles were not non-specific shedding of host membrane but rather conduits for the selective release of mature PV particles. Similar results were also obtained from human rhinovirus 2-infected cells where PS vesicles were found to contain mature viral particles (Figure 3F).

#### Infection by PV Particles in Vesicles Is Dependent on Both the Virus-Specific Receptor and the PS Lipids

We next compared the infection capability of PV particles within vesicles released from cells compared to free viral particles. Free viral particles were obtained by three cycles of quick freeze-thaw of the vesicle fraction. Freeze-thaw does not significantly impact PV infectivity (Strazynski et al., 2002), and no difference in infectivity was observed by plaque assay between equivalent numbers of PV particles isolated by freeze thaw or collection from the supernatant of post-vesicle enrichment fractions (data



**Figure 3. Mature PV, CVB3, and Rhinovirus Particles Are Released in Phosphatidylserine Lipid Vesicles**

(A) Cells infected with PV, CVB3, or rhinovirus were incubated with Alexa 568-Annexin V and imaged by confocal/DIC microscopy at 7 hr p.i. Scale bar represents 5  $\mu$ m for PV and rhinovirus, 10  $\mu$ m for CVB3.

(B) Mock and apoptotic HeLa cells labeled with Alexa 568-Annexin V and imaged by confocal/DIC microscopy. Scale bar represents 10  $\mu$ m.

(C) Dynamics of Alexa 568-Annexin V-labeled PS vesicle release from plasma membrane projections of PV-infected cells at 7 hr p.i.

(D) Quantification of PS vesicles released from enterovirus-infected cells.

(E) VP2/VP0 ratio of whole cell lysate and PS vesicles in PV-infected cells. Annexin V-isolated PS vesicles from PV-infected cells at 7 hr p.i. were analyzed by SDS-PAGE/western with anti-VP2 antibody.

(F) PS vesicles from human rhinovirus-infected cells contain mature rhinoviral particles. Isolated PS vesicles were processed for SDS-PAGE/western analysis with anti-HRV2/VP2 (neutralizing) antibodies.

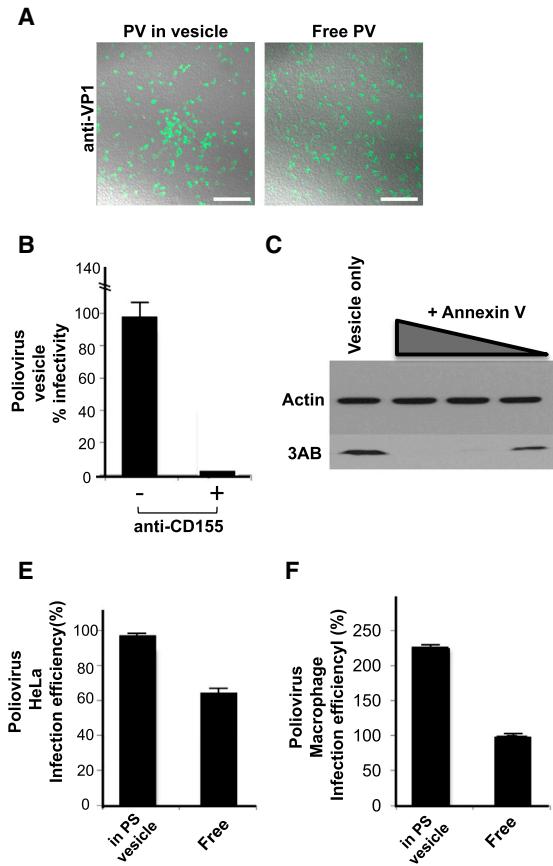
See also Figure S2 and Movie S1.

not shown). The collected extracellular vesicles from PV-infected cells or free PV particles were then incubated with a confluent layer of HeLa cells (Figure 4A). After 4 hr of infection with either vesicles or free viral particles, numerous PV-replicating (based on immunofluorescence labeling of VP1) infected cells were found in both cell populations (Figure 4A).

To determine if the PV particles within vesicles still required the PV receptor CD155 on the host cell for infection, cells were incubated with CD155 neutralizing antibodies prior to exposure to vesicles and viral 3AB replication protein levels were measured after 4 hr of infection. In the presence of neutralizing antibodies, vesicle infectivity was inhibited by >95% indicating that PV ves-

icles were not just simply “fusing” with cells but that the viral particles within the vesicles still required binding to PV receptor for transfer of viral RNA into the host cell cytoplasm (Figure 4B).

We next determined if the infection was dependent on PS lipids of the vesicles. Vesicles isolated by differential centrifugation were incubated with different amounts of Annexin V protein, which binds and masks the PS head-groups on the lipids (Swairjo et al., 1995). After removing any unbound Annexin V, the vesicles were added to HeLa cells and replication measured after 4 hr. Strikingly, masking of the vesicle-associated PS lipids by Annexin V, inhibited PV infection of the host cells in a dose-dependent manner (Figures 4C and 4D). These data indicated



**Figure 4. Infection by PV Particles in Vesicles Is More Efficient Than Free Viral Particles and Is Dependent on Both the Poliovirus Receptor and PS Lipids**

(A) Free or vesicle-associated PV particles were incubated with new recipient cells and replication was detected at 4 hr p.i. by immunolabeling with anti-VP1 antibodies. Scale bar represents 500  $\mu$ m.

(B) CD155/PVR neutralizing antibodies block infection PV particles in vesicles.

(C) Blocking PS lipids on vesicles containing PV particles block infection. Vesicles collected by differential centrifugation were incubated with different amounts of Annexin V protein prior to incubation with recipient HeLa cells. Replication was measured at 4 hr p.i., by SDS-PAGE/western analysis with anti-3AB antibody.

(D) Quantification of western results in (C).

(E) HeLa cells were incubated with equal titers of free and vesicle-associated PV particles. Infection efficiency was determined by quantifying viral 3AB protein levels at peak replication time (4 hr p.i.). Data represented as mean  $\pm$  SD.

(F) Primary human macrophages were incubated with equal titers of free and vesicle-associated PV particles. Infection efficiency was determined by quantifying viral 3AB protein levels at peak replication time (8 hr p.i.). Data represented as mean  $\pm$  SD.

that PS lipids are cofactors for PV infection and that mature infectious PV particles are predominantly in the PS vesicle fraction of vesicles collected by differential centrifugation.

#### Infection by PV Particles in Vesicles Is More Efficient Than Free PV Particles

We next measured and compared the level of infection of host cells when infected with equivalent numbers of PV particles either in vesicles or as free. As a measure of infection efficiency, we quantified and plotted the levels of viral 3AB replication proteins at peak replication times (Figures 4E and 4F). Due to viral RNA synthesis feeding back on viral RNA translation (and vice versa), 3AB levels reflect viral RNA levels (Hsu et al., 2010). We found that at 4 hr p.i. of HeLa cells, viral 3AB levels were ~40% greater in cells incubated with PV particles in vesicles than free PV particles (Figure 4E). This difference was even more striking when primary human macrophages, cells that are specialized to recognize and take up PS lipid-containing cells and vesicles (Fadok et al., 1992), were used as the recipient host. Here vesicle-enclosed PV particles were almost 2-fold greater in infection efficiency compared to free PV particles (Figure 4F).

#### Unilamellar PS Vesicles Released from PV-Infected Cells Contain Multiple Viral Particles

To quantify the clustering of viral particles within vesicles, we collected free and vesicle-associated PV particles and immuno-

labeled them with A12 antibodies. We then imaged the cover-glass-immobilized PV particles by total-internal reflection fluorescence (TIRF) and by DIC microscopy (Figure 5A). Viruses were deposited on coverslip in two formats: vesicle-free fraction ("Free"), or vesicle-embedded fraction ("Vesicular"). Imaging revealed cluster of viruses within vesicles, while free viruses yielded a more diffuse distribution (Figure 5A). We quantified the difference in distribution of measured fluorescence by computing the radial autocorrelation functions  $g(r)$  for the intensity map  $I(\vec{u})$ :

$$g(r) = \langle \phi_{\|\vec{v}\| = r} I(\vec{u}) I(\vec{u} + \vec{v}) d\vec{u} \rangle_{\vec{u}}.$$

The observed exponential decay of these autocorrelation functions enabled us to quantify the length scales that characterize the imaged aggregates (Figure 5B). We found that vesicle-embedded viruses yielded a characteristic clustering scale of  $2.0 \pm 0.1 \mu\text{m}$  ( $n = 4$  images), while free virus yielded a five-time smaller characteristic scale of  $0.4 \pm 0.1 \mu\text{m}$  ( $n = 2$  images). However, these estimates are too close to the spatial resolution of TIRF imaging and thus only qualitative. To circumvent this diffraction limit, we then applied the direct stochastic optical reconstruction microscopy (dSTORM) methodology to achieve super-resolution of these viral particles (Figure 5C). dSTORM relies on sequential imaging and fitting to achieve a spatial resolution of individual dyes within 30 nm (Bates et al., 2007; Baddeley et al., 2009; Dempsey et al., 2011).

In order to assess the degree of clustering within these viral spreads we used the Ripley's K statistical test. By definition,

$$K(r) = \frac{A}{n^2} \sum_{i \neq j} \delta(d_{ij}/r), \text{ with } \delta(x) = \begin{cases} 1 & \text{for } x \leq 1 \\ 0 & \text{for } x > 0 \end{cases},$$

with  $d_{ij}$  being the distance between the  $i^{\text{th}}$  and  $j^{\text{th}}$  points,  $A$  is the image area. By definition,  $K(r)/\pi r^2$  is around 1 for a homogenous distribution of points, and larger than 1 for clustered spatial distribution (Veatch et al., 2012; Termini et al., 2014). Based on the electron microscopy images (Figures 2A, 2B, and S2A), we anticipated clusters of viruses within vesicles of 200–400 nm diameter. Hence, we calculated  $K(r = 200 \text{ nm})$  from the dSTORM data to test whether vesicles contained clusters of viruses: we found that  $K(r = 200 \text{ nm}) = 2.4 \pm 0.9$  for free viruses and  $K(r = 200 \text{ nm}) = 13.7 \pm 6.7$  for viruses within vesicles (Figure 5D). Hence, super-resolution microscopy does confirm that these viral vesicles do pack large numbers of viruses.

Consistent with these findings, transmission electron micrographs of PS vesicles (after isolation by Annexin V-coupled microbeads) revealed multiple clustered viral particles surrounded by a single membrane bilayer (Figure 5E). Per 200 nm cross-section, a single vesicle contained on average  $19 \pm 3$  PV particles ( $n = 6$  vesicles).

The presence of a single bilayer, as opposed to multiple bilayers, also indicated that the double-membraned autophagosome-like organelles had fused with the plasma membrane, rather than budded, in order to release PS vesicles. Thus, the PS lipids on the extracellular membrane leaflet of the vesicle are in a compartment that is topologically equivalent to the luminal membrane leaflet of the double-membraned autophagosome-like organelle. The ER is a major membrane source for autophagosomes (Hamasaki et al., 2013) and has luminal membrane leaflet enriched in PS lipids (Kay et al., 2012). Given that the isolated PS vesicles also contain ER-resident proteins (Figure S2B), it is highly likely that the autophagosome-like organelles, and thereby the released PS vesicles, originated from ER and/or ER-derived replication organelle membranes.

### Vesicles Allow Multiple Viral RNA Molecules to Be Collectively Transferred into Cells

Given our dSTORM and TEM finding, we conjectured that multiple PV particles within a single vesicle might allow multiple viral genomes to be simultaneously transferred into a single cell. To test this hypothesis, we performed single molecule fluorescence *in situ* hybridization (single molecule RNA FISH) (Raj et al., 2008). Multiple oligonucleotide fluorescent probes hybridizing to an RNA molecule allow sufficient sensitivity to detect single RNA molecules within cells (Shaffer et al., 2013; Lubeck and Cai 2012).

Forty-eight fluorescently labeled nucleic acid probes, each 22 nucleotides in length, complementary to the PV genome, were synthesized. Fixed numbers of HeLa cells were incubated with different titers of either free PV particles or vesicle-associated PV particles for 1.5 hr and processed for FISH labeling as described previously (Shaffer et al., 2013) (Figure 5F). Note that we could exclude the possibility of any viral RNA synthesis occurring during this incubation time because cells with and without

Guanidine HCl, an inhibitor of viral RNA synthesis, showed similar amounts of viral RNA molecules per cell (Figures 5G and S3).

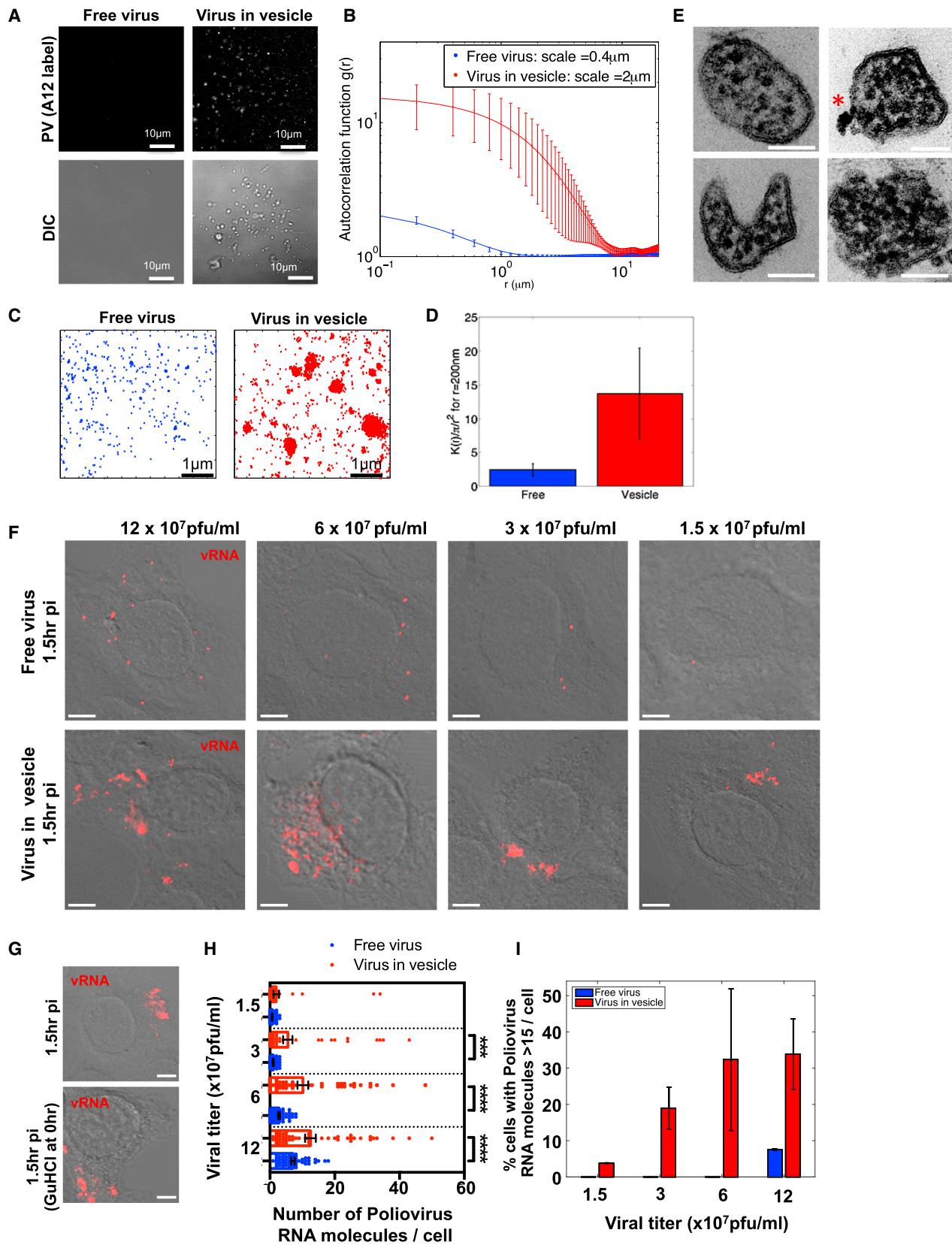
For cells incubated with free PV particles, discrete fluorescent puncta, spatially segregated from one another in the cytoplasm were detected in individual cells (Figure 5F, free virus). In contrast for cells incubated with equivalent titers of PV particles in vesicles, there were many more fluorescent puncta, spatially juxtaposed (Figure 5F, virus in vesicle). The lowest multiplicity of infection with free viral particles was observed at  $1.5 \times 10^7$  pfu/ml where there was on average only one fluorescent puncta per cell (Figures 5F and 5H, free virus). Hence, the size of these puncta was approximated as a single viral RNA molecule and used in subsequent quantification and analysis of FISH data. At each titer, for either free virus or vesicle-associated virus, 55 cells were randomly chosen and viral RNA molecules counted in each cell. The number of viral RNA molecules per cell were subsequently plotted in Figure 5H.

From our quantification, we found that at  $12 \times 10^7$  pfu/ml and  $6 \times 10^7$  pfu/ml, there were ~40% and ~75%, respectively, more viral RNA molecules per cell when PV particles were presented in vesicles than as free virus (Figure 5H). Indeed, for cells incubated with free PV particles, we found that as the viral titer decreased 4-fold (from  $12 \times 10^7$  pfu/ml to  $3 \times 10^7$  pfu/ml), there was a ~90% decrease in the number of PV RNA molecules per cell whereas for cells incubated with PV particles in vesicles, this decrease was significantly less, only ~40% (Figure 5H). Furthermore, at any given titer there were significantly more cells with  $>15$  PV RNA molecules within them when they had been infected with PV in vesicles as opposed to free PV (Figure 5I). These data are consistent with our dSTORM findings (Figures 5A–5D) and transmission electron micrographs of isolated PS vesicles (Figure 5E) and indicate that vesicles contain multiple PV particles, which enable multiple viral RNA genomes to be transferred en bloc into a cell. Note that there was a large variation in the number of viral RNA genomes per cell when cells were infected with PV particles in vesicles as opposed to free PV particles (Figures 5H and 5I). Since de novo viral RNA synthesis can be ruled out, this variation is likely due to differences in the number of viral particles packaged per vesicle as well as contamination from free particles due to possible vesicle lysis.

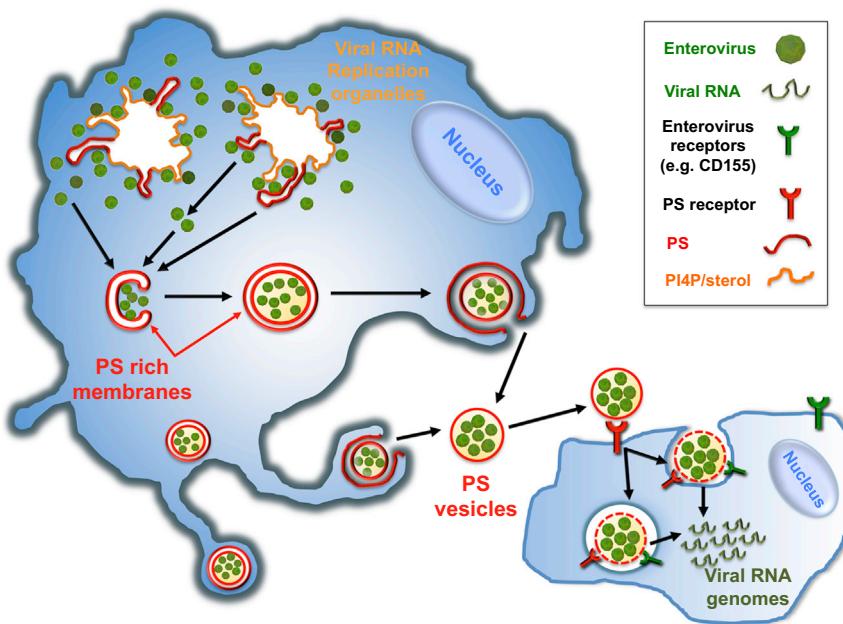
### DISCUSSION

Here, we have shown that multiple mature infectious enteroviral particles are released in single unilamellar PS lipid vesicles, which in turn enables multiple viral genomes to be collectively transferred to an individual cell in a new round of infection. These PS vesicles containing multiple viral genomes appear to be significantly more efficient in infection and may facilitate genetic cooperation among viral genomes.

We first detected assembled PV capsids at replication organelles where viral RNA was synthesized (Figures 1A and 6). This close juxtaposition of capsids with viral RNA synthesis sites would serve to immediately encapsidate the viral RNA and thereby facilitate efficient genome packaging as well as protection of viral genomes from host defenses. These viral particles then translocated into the cytoplasm, which was not only temporally coincident with a cessation of viral RNA synthesis but also



(legend on next page)



could be prematurely triggered by inhibiting RNA synthesis with GuHCl (Figure S1A). Enteroviral 2C proteins may modulate the close coupling between RNA synthesis kinetics and capsid release from replication organelles as 2C proteins are not only localized to the replication organelles and required for viral RNA synthesis but also physically interact with capsids (Liu et al., 2010).

Once the capsids dissociated from the replication organelles, they were sequestered within double-membraned LC3-II-positive autophagosome-like organelles. How the viral cargo is recognized and captured within these organelles, the capsid-associated determinants, and whether LC3-II proteins play a role in these processes as they do in canonical autophagy pathways (Rogov et al., 2014) is currently unknown. Interestingly, unlike canonical autophagosomes that fuse with lysosomes and degrade their cytoplasmic cargo, these capsid-containing or-

ganelles fused with the plasma membrane to non-lytically release >80% of the PV particles into the extracellular environment within unilamellar vesicles of size range 200–400 nm. Notably, the SNARE protein syntaxin 17 was not localized to the autophagosome-like organelles to regulate their fusion with lysosomes, but instead was found sequestered away at the replication organelles (Figure S1I). Whether specific enteroviral proteins actively modulate its subcellular localization or it is an indirect consequence of the affinity of the syntaxin 17 hydrophobic hairpin tail for the PI4P/cholesterol rich replication organelle membranes remains to be investigated. This type of non-conventional secretion of autophagosomal membranes from the cell has never been reported and identifying the machinery regulating this process, including determining which cytoskeletal components and SNARE proteins are utilized for movement out to the periphery and fusion with the plasma membrane, may provide novel therapeutic targets to block enterovirus release from cells.

**Figure 5. PS Vesicles Contain Clustered PV Particles, which Enable Multiple Viral RNA Genomes to Be Collectively Transferred into a New Host Cell**

- (A) TIRF and DIC images of free and vesicle-associated PV particles labeled with Atto488-labeled A12 antibody.
- (B) Difference in distribution of fluorescence in (A) by computation of the radial autocorrelation function  $g(r)$ .
- (C) dSTORM imaging free and vesicle-associated viral particles labeled with Atto488-A12 antibody.
- (D) Calculation of Ripley's  $K$  function to assess the degree of clustering of vesicle-associated PV particles relative to free particles. Data represented as mean  $\pm$  SEM.
- (E) PS vesicles isolated from PV-infected cells using Annexin V microbeads were imaged by transmission electron microscopy. Note the numerous electron-dense viral particles in each vesicle and the unilamellar surrounding membrane. Asterisk shows Annexin V microbead attached to the exterior of one vesicle. Scale bars represent 100 nm.
- (F) Collected intact vesicles or free viral particles from PV-infected cells were incubated with a confluent layer of new host cells at different viral titers. Viral RNAs were monitored by single molecule RNA FISH and imaged by dual confocal/DIC microscopy at 1.5 hr p.i. Shown are images of single HeLa cells infected with either free viral particles or vesicle-associated viral particles. Images presented were acquired with the same microscopy settings. Scale bar represents 2  $\mu$ m.
- (G) Collected vesicles were incubated with cells with/without GuHCl for 1.5 hr and viral RNA molecules were monitored by single molecule RNA FISH (see also Figure S3 for quantification). Scale bar represents 2  $\mu$ m.
- (H) Quantification of the number of viral RNA molecules per cell in cells infected with either free ( $n = 55$  cells) or vesicle-associated PV particles ( $n = 55$  cells). \*\*\* $p < 2.10^{-3}$ ; \*\*\*\* $p < 10^{-4}$ .
- (I) Percent of cells with 15 or greater PV RNA puncta was quantified for cells infected with either free or vesicle-associated PV particles for each viral titer shown. Data represented as mean  $\pm$  SEM.

**Figure 6. Model**

Assembled mature enteroviruses are released from the replication organelles into the cytoplasm. Clusters of multiple viral particles are selectively captured by double-membraned organelles that originate from the ER and ER-derived replication organelles. These double-membraned organelles, which resemble autophagosomes, contain PS lipids on both the luminal and cytoplasmic leaflets of their membranes. They fuse with the plasma membrane and release a unilamellar PS-lipid-enriched vesicle, containing multiple viral particles, into the extracellular medium. This vesicle then facilitates infection in a PS-lipid and viral receptor-dependent mechanism resulting in the collective transfer to a new recipient host cell of multiple viral RNA genomes. This mode of viral transmission enhances infection efficiency and potentially allows for genetic complementation among quasispecies.

PS vesicles non-lytically released from cells were selectively enriched in mature enteroviral particles (Figure 3). While we cannot exclude the possibility that non-PS vesicles may also contribute to the non-lytic release of enteroviral particles, the observed significant inhibition of subsequent infection when PS is blocked (Figures 4C and 4D) suggest that PS vesicles constitute a large fraction of the non-lytic conduit for enteroviral release from cells. Live-cell time-lapse imaging in the presence of fluorescently labeled Annexin V protein revealed that the outer membrane leaflet of the released vesicle and hence topologically equivalent to the luminal membrane leaflet of the double membranous organelle pre-fusion, contained PS lipids (Figures 3A–3D; Movie S1). The ER and/or the ER-derived replication organelles both have PS lipids on their luminal and cytoplasmic leaflets (Leventis and Grinstein, 2010; Lev, 2012) (Figure 1D), and moreover, the PS can flip between these two leaflets (Clark, 2011). Thus it is likely that the host source for the autophagosome-like organelles is the ER and/or ER-derived replication organelles. Supporting this conclusion the released PS vesicles indeed contained ER markers including the integral ER membrane protein calnexin (Figure S2B).

Super-resolution imaging and transmission electron microscopy revealed extracellular PS vesicles to be containing multiple viral particles, at least 20 per 200 nm cross-section (Figures 5A–5E). Supporting these data, in single molecule RNA FISH experiments, we found that infection by PV particles in vesicles allowed the collective transfer of multiple viral genomes into a single host cell (Figures 5F–5I). One important implication of these findings is that it ties the fate of individual viral genomes from previous rounds of replication to each other and thereby may provide selective advantages in terms of replication kinetics and genetic diversity relative to free viral particle genomes (i.e., not in vesicle). Indeed, infection efficiency was significantly higher when cells were infected with PV particles in vesicles as opposed to an equivalent number of free virus particles (Figures 4E and 4F). Enteroviruses as well as all other positive-stranded RNA viruses, are enormously diverse in genomic variety due to the inherent error rates and lack of proof reading in their RNA polymerases that can generate large numbers of viral quasispecies after even a single round of infection (Bordería et al., 2011). Vesicular transfer of multiple particles among cells would increase the chances of genetic complementation among viral quasispecies, potentially benefiting the replication efficiency of otherwise attenuated or weak genomes and enabling them to maintain a presence in the genetic pool. Indeed our findings may provide a cellular mechanism to explain the results of Vignuzzi et al. (2006) where cooperative interactions between neurotropic and non-neurotropic PV quasispecies were reported. Second, when multiple viral genomes are transferred to a single cell, the likelihood of one or more genomes surviving the hostile host environment to override host defenses and replicate may be higher. Third, for positive-stranded RNA viruses, rather than a single genome having to switch between RNA translation and RNA synthesis activities (until sufficient levels of viral RNA have been synthesized to partition those functions among genomes), multiple genomes could right from the start of infection partition RNA translation and RNA synthesis functions among themselves and enhance overall replication kinetics

and viral protein levels. Finally the PS lipids on the vesicles themselves could enhance infection efficiency by attracting PS scavenging cells, such as macrophages and dendritic cells, to take up the viral genomes and provide a host environment for replication. In particular, we found that PV particles within vesicles could replicate significantly more efficiently within primary macrophages than free PV particles (Figure 4F).

What is the mechanism whereby viral particles within vesicles infect host cells? We found infection to be dependent not only on the virus-specific receptor expressed by the host but also on the PS lipids associated with the vesicle having access to the host cell (Figures 4C and 4D). One potential mechanism to explain these findings is that PS lipids on vesicles engage PS receptors on the recipient host cell prior to the viral particles engaging their own receptors. The binding to PS receptors can trigger phagocytic uptake of vesicles (Hoffmann et al., 2001) followed by lysis or permeabilization of the vesicle within the endosomal compartment to then enable viral particles to engage their specific receptors (Figure 6). Equally possible, binding to PS receptors may lead to permeabilization of the vesicle on the cell surface, resulting in the release of a concentrated bolus of viral particles in the immediate vicinity of the cell. The latter mechanism may also provide access to neutralizing antibodies which bind capsids and block infection. Reliance on PS lipids and PS receptors for infection has been documented for a number of other viruses including vaccinia, Dengue, Ebola, hepatitis A virus, and HIV (Sui et al., 2006; Mercer and Helenius, 2009; Feng et al., 2014; Morizano and Chen, 2014). The specific use of PS lipids by enteroviruses as well as other viruses to traffic between cells may have significant in vivo implications for viral pathogenesis and tissue tropism. In particular, the infection of primary macrophages, the major PS sensing cells in the body, may provide enteroviruses with the ability to target a key cell subset of the immune system while suppressing its inflammatory responses, as PS lipids have been well documented to inhibit inflammatory cytokine production by macrophages (Hochreiter-Hufford and Ravichandran, 2013). PS-lipid vesicles may also help enteroviruses exploit the natural motility of macrophages and help spread them distant sites including perhaps the CNS (Ousman and Kubes, 2012).

In summary, we report here a novel mode of viral transmission among cells where multiple viral particles are clustered and collectively released within PS-lipid-enriched vesicles. This provides greater infection efficiency and potentially an opportunity for cooperation and complementation among viral quasispecies. This mode of transmission links the fate of multiple viral particles to one another and may have implications for maintaining viral genetic diversity within viral evolution.

## EXPERIMENTAL PROCEDURES

All detailed protocols and information regarding plasmids, antibodies, cell culture, virus infection, and propagation are provided in the [Extended Experimental Procedures](#).

### Immunofluorescence

Cells were plated on glass coverslips and fixed with 4% PFA for 15 min at RT. Cells were permeabilized with either 0.2% Saponin or 0.1% Triton X-100 and sequentially incubated with primary and fluorophore-tagged secondary

antibodies. Coverslips were mounted in Fluoromount-G (Southern Biotech) and imaged.

### Confocal Microscopy

All confocal images were obtained with an LSM780 laser scanning confocal microscope system (Carl Zeiss) and images were analyzed with either Zen (Carl Zeiss) or Image J (NIH) software.

### dSTORM

Free PV particles and PV particles in vesicles were plated on gridded glass bottom dishes and fixed with 4% PFA for 15 min at room temperature. Subsequently they were permeabilized with 0.2% saponin and incubated with Atto488-conjugated A12 antibodies. An oxygen-scavenging PBS solution (10 mM NaCl, 0.5 mg/ml glucose oxidase, 40 g/ml catalase, 2% glucose, and 10 mM MEA) was used for imaging. dSTORM images were obtained on a Zeiss ELYRA PS.1 system (Carl Zeiss). Images were acquired with a Plan-Apochromat 100 $\times$ /1.46 oil immersion objective and an Andor iXon 885 EMCCD camera. A total of 20,000 images were acquired per sample with an exposure time of 33 ms. Raw images were reconstructed and analyzed with ZEN software (Carl Zeiss) and MATLAB (MathWorks) using methodology from Veatch et al. (2012) and Termini et al. (2014).

### Structured Illumination Microscopy

Super-resolution 3D-structured illumination microscopy (SIM) imaging was performed on a Zeiss ELYRA PS.1 system (Carl Zeiss). Images were acquired with a Plan-Apochromat 63 $\times$ /1.40 oil immersion objective and an Andor iXon 885 EMCCD camera. Fifteen images per plane (five phases, three rotations) and 0.125 mm z section of 3 mm height were required for generating super resolution images. Raw images were reconstructed and processed to demonstrate structure with greater resolution by the ZEN 2011 software (Carl Zeiss).

### Single Molecule RNA FISH

We performed single molecule RNA FISH according to Shaffer et al. (2013). Cells were incubated with either free or vesicle-associated PV particles for 1.5 hr. Cells were subsequently fixed in pre-chilled methanol ( $-20^{\circ}\text{C}$ ) for 10 min. The methanol was removed and the cells were hybridized with 10  $\mu\text{l}$  of hybridization buffer containing 4  $\mu\text{M}$  PV probe, 10% formamide, 2 $\times$  SSC, and 10% dextran sulfate for 10 min at  $37^{\circ}\text{C}$ . Next, the samples were washed three times with pre-warmed wash buffer (10% formamide and 2 $\times$  SSC) at  $37^{\circ}\text{C}$  for 1 min and imaged with Zeiss LSM780 confocal microscope.

### Annexin V Labeling

Cells for live imaging were grown on coverslip-bottomed Lab-Tek chambers (Thermo Fisher) and infected with PV for 7 hr. Cells were then replaced in imaging media (DMEM Phenol Red free supplemented with 10% FBS and 50 mM HEPES [pH 7.3]). Alexa 568-Annexin V was added on the cells and imaging was performed on a Zeiss LSM780 Confocal Laser Scanning microscope (Carl Zeiss) equipped with 458 nm, 488 nm, 514 nm, 565 nm, and 633 nm laser lines and detecting system for fluorescence and DIC imaging. The microscope was additionally equipped with a heating stage and incubator with temperature, humidity, and  $\text{CO}_2$  control for live-cell imaging.

### SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, three figures, and one movie and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2015.01.032>.

### ACKNOWLEDGMENTS

The authors would like to thank Yasmine Belkaid, Konstantin Chumakov, Ana Maria Cuervo, James Hogle, Gerald Feldman, Jennifer Jones, Jennifer Lippincott-Schwartz, Sanford Simon, Radek Dobrowski, Wilma Friedman, Ellie Ehrenfeld and members of the Altan-Bonnet lab for fruitful discussions. The authors would like to especially thank Wen-Chin Tseng for help in constructing the graphical cartoons. The authors also thank Frank Macaluso and Geoffrey

Perumal of the Albert Einstein College of Medicine Analytical Imaging facility for technical support. A.C., P.T., and C.P. were supported by grant AI091985-01A1; G.A.B. was supported by grant AI083408; H.M., P.J., and E.W. were supported by grant AI15122; and Y.C., W.D., M.H., C.B., P.C., E.S., and N.A.B. were supported by intramural NIH funds.

Received: October 2, 2014

Revised: December 13, 2014

Accepted: January 12, 2015

Published: February 12, 2015

### REFERENCES

- Baddeley, D., Jayasinghe, I.D., Cremer, C., Cannell, M.B., and Soeller, C. (2009). Light-induced dark states of organic fluorophores enable 30 nm resolution imaging in standard media. *Biophys. J.* **96**, L22–L24.
- Barton, D.J., and Flanagan, J.B. (1997). Synchronous replication of poliovirus RNA: initiation of negative-strand RNA synthesis requires the guanidine-inhibited activity of protein 2C. *J. Virol.* **71**, 8482–8489.
- Bates, M., Huang, B., Dempsey, G.T., and Zhuang, X. (2007). Multicolor super-resolution imaging with photo-switchable fluorescent probes. *Science* **317**, 1749–1753.
- Bird, S.W., Maynard, N.D., Covert, M.W., and Kirkegaard, K. (2014). Nonlytic viral spread enhanced by autophagy components. *Proc. Natl. Acad. Sci. USA* **111**, 13081–13086.
- Bordería, A.V., Stapleford, K.A., and Vignuzzi, M. (2011). RNA virus population diversity: implications for inter-species transmission. *Curr. Opin. Virol.* **1**, 643–648.
- Brandenburg, B., and Zhuang, X. (2007). Virus trafficking - learning from single-virus tracking. *Nat. Rev. Microbiol.* **5**, 197–208.
- Brandenburg, B., Lee, L.Y., Lakadamyali, M., Rust, M.J., Zhuang, X., and Hogle, J.M. (2007). Imaging poliovirus entry in live cells. *PLoS Biol.* **5**, e183.
- Chen, Z., Chumakov, K., Dragunsky, E., Kouikava, D., Makiya, M., Neverov, A., Rezapkin, G., Sebrell, A., and Purcell, R. (2011). Chimpanzee-human monoclonal antibodies for treatment of chronic poliovirus excretors and emergency postexposure prophylaxis. *J. Virol.* **85**, 4354–4362.
- Chen, Z., Fischer, E.R., Kouikava, D., Hansen, B.T., Luttko, S.J., Bidzheva, B., Makiya, M., Agullo, L., Purcell, R.H., and Chumakov, K. (2013). Cross-neutralizing human anti-poliovirus antibodies bind the recognition site for cellular receptor. *Proc. Natl. Acad. Sci. USA* **110**, 20242–20247.
- Clark, M.R. (2011). Flippin' lipids. *Nat. Immunol.* **12**, 373–375.
- Dempsey, G.T., Vaughan, J.C., Chen, K.H., Bates, M., and Zhuang, X. (2011). Evaluation of fluorophores for optimal performance in localization-based super-resolution imaging. *Nat. Methods* **8**, 1027–1036.
- den Boon, J.A., and Ahlquist, P. (2010). Organelle-like membrane compartmentalization of positive-strand RNA virus replication factories. *Annu. Rev. Microbiol.* **64**, 241–256.
- Doceul, V., Hollinshead, M., van der Linden, L., and Smith, G.L. (2010). Repulsion of superinfecting virions: a mechanism for rapid virus spread. *Science* **327**, 873–876.
- Ehrenfeld, E., Maizel, J.V., and Summers, D.F. (1970). Soluble RNA polymerase complex from poliovirus-infected HeLa cells. *Virology* **40**, 840–846.
- Fadok, V.A., Voelker, D.R., Campbell, P.A., Cohen, J.J., Bratton, D.L., and Henson, P.M. (1992). Exposure of phosphatidylserine on the surface of apoptotic lymphocytes triggers specific recognition and removal by macrophages. *J. Immunol.* **148**, 2207–2216.
- Feng, Z., Hensley, L., McKnight, K.L., Hu, F., Madden, V., Ping, L., Jeong, S.H., Walker, C., Lanford, R.E., and Lemon, S.M. (2013). A pathogenic picornavirus acquires an envelope by hijacking cellular membranes. *Nature* **496**, 367–371.
- Feng, Z., Li, Y., McKnight, K.L., Hensley, L., Lanford, R.E., Walker, C.M., and Lemon, S.M. (2014). Human pDCs preferentially sense enveloped hepatitis A viruses. *J. Clin. Invest.* **125**, 169–176.
- Flint, S.J., Enquist, L.W., Racaniello, V.R., and Skalka, A.M. (2009). Principles of Virology, Third Edition (ASM Press).

Hamasaki, M., Furuta, N., Matsuda, A., Nezu, A., Yamamoto, A., Fujita, N., Oomori, H., Noda, T., Haraguchi, T., Hiraoka, Y., et al. (2013). Autophagosomes form at ER-mitochondria contact sites. *Nature* 495, 389–393.

Hochreiter-Hufford, A., and Ravichandran, K.S. (2013). Clearing the dead: apoptotic cell sensing, recognition, engulfment, and digestion. *Cold Spring Harb. Perspect. Biol.* 5, a008748.

Hoffmann, P.R., deCathelineau, A.M., Ogden, C.A., Leverrier, Y., Bratton, D.L., Daleke, D.L., Ridley, A.J., Fadok, V.A., and Henson, P.M. (2001). Phosphatidylserine (PS) induces PS receptor-mediated macropinocytosis and promotes clearance of apoptotic cells. *J. Cell Biol.* 155, 649–659.

Hsu, N.Y., Illytska, O., Belov, G., Santiana, M., Chen, Y.H., Takvorian, P.M., Pau, C., van der Schaar, H., Kaushik-Basu, N., Balla, T., et al. (2010). Viral reorganization of the secretory pathway generates distinct organelles for RNA replication. *Cell* 141, 799–811.

Illytska, O., Santiana, M., Hsu, N.Y., Du, W.L., Chen, Y.H., Viktorova, E.G., Belov, G., Brinker, A., Storch, J., Moore, C., et al. (2013). Enteroviruses harness the cellular endocytic machinery to remodel the host cell cholesterol landscape for effective viral replication. *Cell Host Microbe* 14, 281–293.

Itakura, E., Kishi-Itakura, C., and Mizushima, N. (2012). The hairpin-type tail-anchored SNARE syntaxin 17 targets to autophagosomes for fusion with endosomes/lysosomes. *Cell* 151, 1256–1269.

Jackson, W.T., Giddings, T.H., Jr., Taylor, M.P., Mulinyawe, S., Rabinovitch, M., Kopito, R.R., and Kirkegaard, K. (2005). Subversion of cellular autophagosomal machinery by RNA viruses. *PLoS Biol.* 3, e156.

Kay, J.G., Koivusalo, M., Ma, X., Wohland, T., and Grinstein, S. (2012). Phosphatidylserine dynamics in cellular membranes. *Mol. Biol. Cell* 23, 2198–2212.

Kirkegaard, K., and Jackson, W.T. (2005). Topology of double-membraned vesicles and the opportunity for non-lytic release of cytoplasm. *Autophagy* 1, 182–184.

Koopman, G., Reutelingsperger, C.P., Kuijten, G.A., Keehnen, R.M., Pals, S.T., and van Oers, M.H. (1994). Annexin V for flow cytometric detection of phosphatidylserine expression on B cells undergoing apoptosis. *Blood* 84, 1415–1420.

Lev, S. (2012). Non-vesicular lipid transfer from the ER. *Cold Spring Harb. Perspect. Biol.* 4, 1–17.

Leventis, P.A., and Grinstein, S. (2010). The distribution and function of phosphatidylserine in cellular membranes. *Annu. Rev. Biophys.* 39, 407–427.

Liu, Y., Wang, C., Mueller, S., Paul, A.V., Wimmer, E., and Jiang, P. (2010). Direct interaction between two viral proteins, the nonstructural protein 2C and the capsid protein VP3, is required for enterovirus morphogenesis. *PLoS Pathog.* 6, e1001066.

Lubeck, E., and Cai, L. (2012). Single-cell systems biology by super-resolution imaging and combinatorial labeling. *Nat. Methods* 9, 743–748.

Mercer, J., and Helenius, A. (2009). Virus entry by macropinocytosis. *Nat. Cell Biol.* 11, 510–520.

Morizano, K., and Chen, I.S. (2014). Role of phosphatidylserine receptors in enveloped virus infection. *J. Virol.* 88, 4275–4290.

Nchoutmboube, J.A., Viktorova, E.G., Scott, A.J., Ford, L.A., Pei, Z., Watkins, P.A., Ernst, R.K., and Belov, G.A. (2013). Increased long chain acyl-CoA synthetase activity and fatty acid import is linked to membrane synthesis for development of picornavirus replication organelles. *PLoS Pathog.* 9, e1003401.

Ousman, S.S., and Kubes, P. (2012). Immune surveillance in the central nervous system. *Nat. Neurosci.* 15, 1096–1101.

Owens, R.J., Limn, C., and Roy, P. (2004). Role of an arbovirus nonstructural protein in cellular pathogenesis and virus release. *J. Virol.* 78, 6649–6656.

Raj, A., van den Bogaard, P., Rifkin, S.A., van Oudenaarden, A., and Tyagi, S. (2008). Imaging individual mRNA molecules using multiple singly labeled probes. *Nat. Methods* 5, 877–879.

Robinson, S.M., Tsueng, G., Sin, J., Mangale, V., Rahawi, S., McIntyre, L.L., Williams, W., Kha, N., Cruz, C., Hancock, B.M., et al. (2014). Coxsackievirus B exits the host cell in shed microvesicles displaying autophagosomal markers. *PLoS Pathog.* 10, e1004045.

Rogov, V., Dötsch, V., Johansen, T., and Kirkin, V. (2014). Interactions between autophagy receptors and ubiquitin-like proteins form the molecular basis for selective autophagy. *Mol. Cell* 53, 167–178.

Shaffer, S.M., Wu, M.T., Levesque, M.J., and Raj, A. (2013). Turbo FISH: a method for rapid single molecule RNA FISH. *PLoS ONE* 8, e75120.

Strauss, M., Levy, H.C., Bostina, M., Filman, D.J., and Hogle, J.M. (2013). RNA transfer from poliovirus 135S particles across membranes is mediated by long umbilical connectors. *J. Virol.* 87, 3903–3914.

Strazynski, M., Krämer, J., and Becker, B. (2002). Thermal inactivation of polio-virus type 1 in water, milk and yoghurt. *Int. J. Food Microbiol.* 74, 73–78.

Sui, L., Zhang, W., Chen, Y., Zheng, Y., Wan, T., Zhang, W., Yang, Y., Fang, G., Mao, J., and Cao, X. (2006). Human membrane protein Tim-3 facilitates hepatitis A virus entry into target cells. *Int. J. Mol. Med.* 17, 1093–1099.

Swairjo, M.A., Concha, N.O., Kaetzel, M.A., Dedman, J.R., and Seaton, B.A. (1995). Ca(2+)-bridging mechanism and phospholipid head group recognition in the membrane-binding protein annexin V. *Nat. Struct. Biol.* 2, 968–974.

Takahashi, M., Yamada, K., Hoshino, Y., Takahashi, H., Ichiyama, K., Tanaka, T., and Okamoto, H. (2008). Monoclonal antibodies raised against the ORF3 protein of hepatitis E virus (HEV) can capture HEV particles in culture supernatant and serum but not those in feces. *Arch. Virol.* 153, 1703–1713.

Taylor, M.P., Burgon, T.B., Kirkegaard, K., and Jackson, W.T. (2009). Role of microtubules in extracellular release of poliovirus. *J. Virol.* 83, 6599–6609.

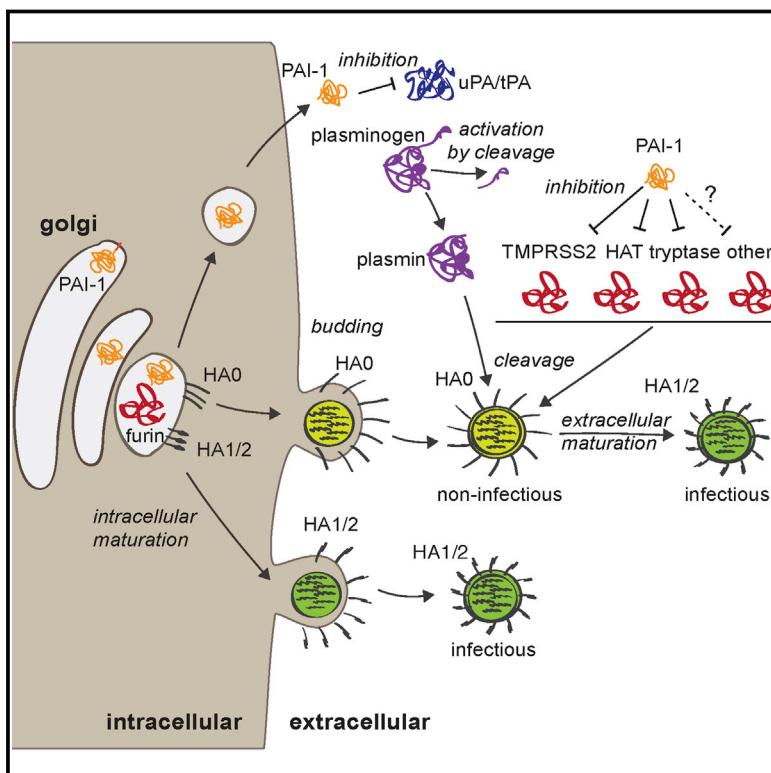
Termini, C.M., Cotter, M.L., Marjon, K.D., Buranda, T., Lidke, K.A., and Gillette, J.M. (2014). The membrane scaffold CD82 regulates cell adhesion by altering  $\alpha$ 4 integrin stability and molecular density. *Mol. Biol. Cell* 25, 1560–1573.

Veatch, S.L., Machta, B.B., Shelby, S.A., Chiang, E.N., Holowka, D.A., and Baird, B.A. (2012). Correlation functions quantify super-resolution images and estimate apparent clustering due to over-counting. *PLoS ONE* 7, e31457.

Vignuzzi, M., Stone, J.K., Arnold, J.J., Cameron, C.E., and Andino, R. (2006). Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. *Nature* 439, 344–348.

# A Serpin Shapes the Extracellular Environment to Prevent Influenza A Virus Maturation

## Graphical Abstract



## Authors

Meike Dittmann,  
Hans-Heinrich Hoffmann, ...,  
Paul D. Bieniasz, Charles M. Rice

## Correspondence

ricec@rockefeller.edu

## In Brief

Plasminogen activator inhibitor (PAI-1) blocks surface glycoprotein maturation of influenza A virus, thus reducing virus spread in the airways and revealing that the innate immune system, driven by type I IFN, uses modulation of the extracellular environment to inhibit viruses.

## Highlights

- *SERPINE1/PAI-1* was identified as an unconventional ISG that acts extracellularly
- PAI-1 inhibits influenza A virus (IAV) spread by inhibiting glycoprotein cleavage
- Endogenous PAI-1 blocks IAV spread in human and murine cells, ex vivo and in vivo
- PAI-1 potentially inhibits other viruses requiring extracellular maturation

# A Serpin Shapes the Extracellular Environment to Prevent Influenza A Virus Maturation

Meike Dittmann,<sup>1</sup> Hans-Heinrich Hoffmann,<sup>1</sup> Margaret A. Scull,<sup>1</sup> Rachel H. Gilmore,<sup>1</sup> Kierstin L. Bell,<sup>1</sup> Michael Ciancanelli,<sup>2</sup> Sam J. Wilson,<sup>3,4</sup> Stefania Crotta,<sup>5</sup> Yingpu Yu,<sup>1</sup> Brenna Flatley,<sup>1</sup> Jing W. Xiao,<sup>1</sup> Jean-Laurent Casanova,<sup>2,6,7,8,9</sup> Andreas Wack,<sup>5</sup> Paul D. Bieniasz,<sup>3</sup> and Charles M. Rice<sup>1,\*</sup>

<sup>1</sup>Laboratory of Virology and Infectious Disease, The Rockefeller University, New York, NY 10065, USA

<sup>2</sup>St. Giles Laboratory of Human Genetics of Infectious Diseases, Rockefeller Branch, The Rockefeller University, New York, NY 10065, USA

<sup>3</sup>Howard Hughes Medical Institute, Laboratory of Retrovirology, Aaron Diamond AIDS Research Center, The Rockefeller University, New York, NY 10016, USA

<sup>4</sup>Institute of Infection, Immunity and Inflammation, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow G12 8TA, UK

<sup>5</sup>Division of Immunoregulation, MRC National Institute for Medical Research, Mill Hill, London NW7 1AA, UK

<sup>6</sup>Howard Hughes Medical Institute, NY 10065, USA

<sup>7</sup>Laboratory of Human Genetics of Infectious Diseases, Necker Branch, Imagine Institute, Inserm, 75015 Paris, France

<sup>8</sup>Paris Descartes University, 75015 Paris, France

<sup>9</sup>Pediatric Hematology-Immunology Unit, Necker Hospital for Sick Children, 75015 Paris, France

\*Correspondence: ricec@rockefeller.edu

<http://dx.doi.org/10.1016/j.cell.2015.01.040>

## SUMMARY

Interferon-stimulated genes (ISGs) act in concert to provide a tight barrier against viruses. Recent studies have shed light on the contribution of individual ISG effectors to the antiviral state, but most have examined those acting on early, intracellular stages of the viral life cycle. Here, we applied an image-based screen to identify ISGs inhibiting late stages of influenza A virus (IAV) infection. We unraveled a directly antiviral function for the gene *SERPINE1*, encoding plasminogen activator inhibitor 1 (PAI-1). By targeting extracellular airway proteases, PAI-1 inhibits IAV glycoprotein cleavage, thereby reducing infectivity of progeny viruses. This was biologically relevant for IAV restriction *in vivo*. Further, partial PAI-1 deficiency, attributable to a polymorphism in human *SERPINE1*, conferred increased susceptibility to IAV *in vitro*. Together, our findings reveal that manipulating the extracellular environment to inhibit the last step in a virus life cycle is an important mechanism of the antiviral response.

## INTRODUCTION

As obligate intracellular parasites, the fate of viruses is intricately linked to the metabolism of their host cells. Viruses exploit host cell mechanisms throughout their life cycle, from entry to replication, assembly and egress, and finally maturation. Each step in this cycle is a potential point for antiviral intervention.

Virus-infected cells produce interferons (IFNs) that trigger transcriptional upregulation of IFN-stimulated genes (ISGs) (reviewed in Schneider et al., 2014). The combined action of hundreds of expressed ISG proteins creates multiple lines of de-

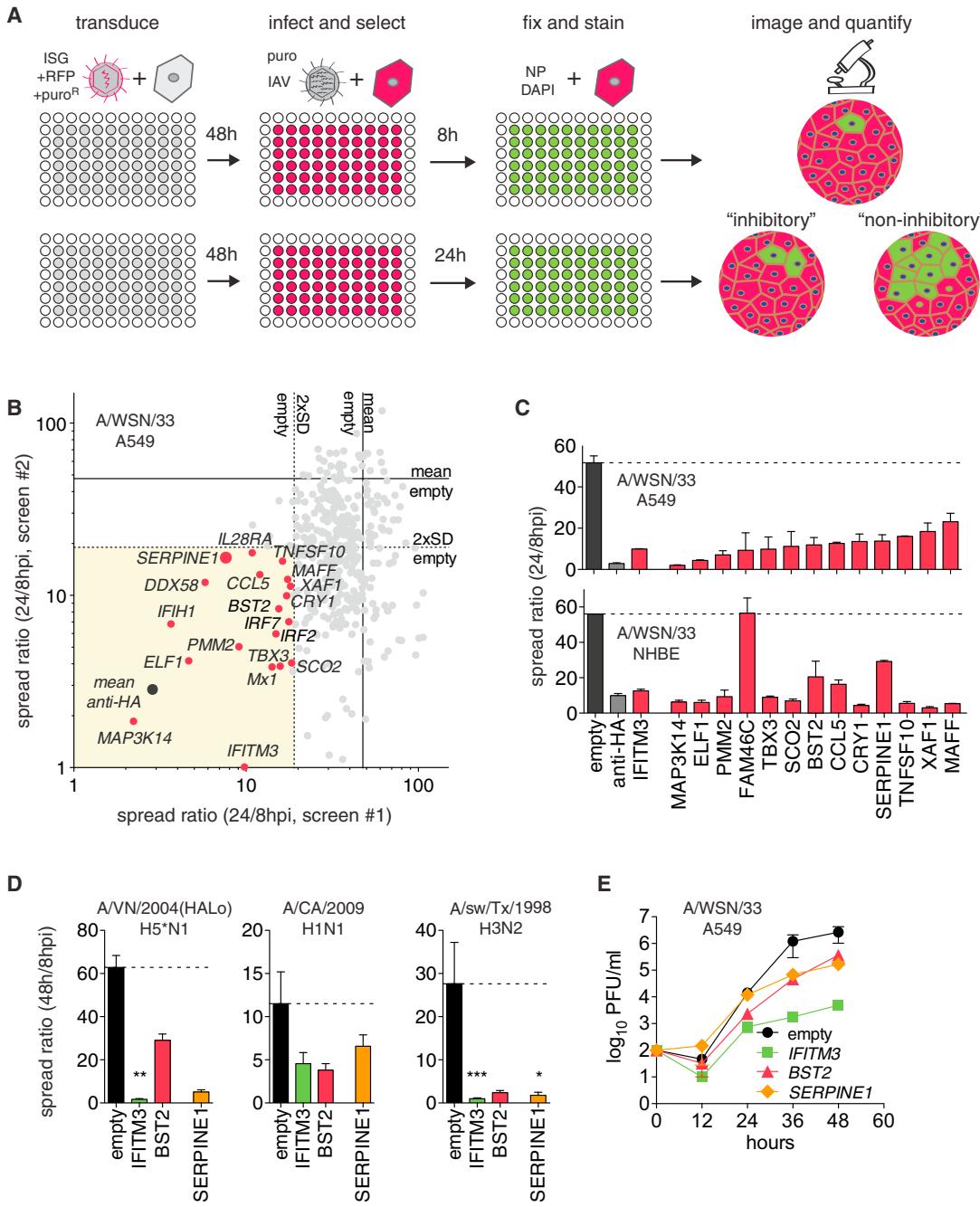
fense against viral infection. It is not surprising that different ISGs block different pathways and have varying potency against different classes of viruses (Schoggins et al., 2014). Our knowledge of ISG effector functions has been somewhat biased toward inhibition of early replication stages, such as entry, or viral translation initiation. However, viral assembly and release have proven to be an Achilles heel for viruses and are the targets of therapeutic small molecules (influenza A virus [IAV] neuraminidase inhibitors; HIV-1 protease inhibitors) and intrinsic ISGs such as tetherin (Neil et al., 2008) or viperin (Wang et al., 2007).

In this study, we established a screen that would also uncover host factors inhibiting late stages of the viral replication cycle, using IAV as a model virus. Applying an approach that accurately monitors several rounds of viral replication in the presence of single overexpressed ISGs, we discovered a direct antiviral function of the well-known gene *SERPINE1*, which encodes PAI-1 (Ny et al., 1986). We demonstrate that PAI-1 inhibits IAV maturation by targeting host proteases needed for viral glycoprotein cleavage and that this is physiologically relevant in natural proteolytic landscapes both *in vitro* and *in vivo*. Finally, we link a human genetic polymorphism in extracellular PAI-1 production to enhanced IAV susceptibility *in vitro*. In all, we show that shaping the extracellular environment can be a powerful mechanism in the vast arsenal of the innate immune response against invading pathogens.

## RESULTS

### A Gain-of-Function Screen Reveals *SERPINE1* Encoding PAI-1 as an Inhibitor of IAV Spread

To identify ISG effectors targeting late stages of the IAV life cycle, we established a high-throughput, image-based screen using an extended version of a recently published ISG library comprising 401 cDNAs (Schoggins et al., 2011; Table S1). Human lung adenocarcinoma cells (A549) were transduced with lentiviral vectors to express individual ISGs (Figure 1A). After 48 hr, cells were challenged with IAV WSN/33 (H1N1) at an MOI of 0.01.



**Figure 1. High-Throughput Microscopy Screens for Inhibitors of IAV Spread**

(A) Screening workflow. Shown are hypothetical effects of expressing inhibitory (antiviral) or non-inhibitory ISGs on single or multiple rounds of virus replication. Red, transduced cells; green, infected cells; blue, DAPI-stained nuclei.

(B) Effect of 401 single ISGs on IAV spread. ISGs inhibiting more than 2-fold SD in two independent screens are shown in red. Spread ratio, the ratio of infected cells at 24/8 hpi. A positive control for inhibition is  $\alpha$ -HA antibody.

(C) Confirmation assays for selected ISGs on A549 cells or primary NHBE cells. Data are represented as mean  $\pm$  SEM from  $n = 6$  values in two independent experiments for A549, and  $n = 3$  for NHBE cells.

(D) SERpine1(PAI-1)-mediated inhibition of A/Vietnam/1203/2004(HALo), A/California/04/2009, and A/sw/Texas/4199-2/1998 in A549 cells. Empty vector, negative control; IFITM3 and BST2 (tetherin), positive controls. Data are represented as mean  $\pm$  SEM from  $n = 4$  independent experiments. One-way ANOVA and Dunn's multiple comparison test versus "empty."

(E) ISG-expressing A549 cells were infected with IAV WSN/33 at MOI 0.01, and virus titers were measured by plaque assay on MDCK cells. Data are represented as mean  $\pm$  SEM from  $n = 4$  independent experiments.

See also Figure S1.

"Spread ratio" was calculated from the number of infected cells at 24 hr post-infection (hpi) relative to 8 hpi for each ISG (Figures 1A and S1A).

The screen was performed twice, using independently generated lentivirus libraries (Figure 1B).  $\alpha$ -HA antibody, with a spread ratio of ~2, was a positive control, whereas empty vector controls had a spread ratio of 50 to 60 (Figures S1C and 1B). Nineteen ISGs reduced the IAV spread ratio to <20, greater than two SDs from the empty vector control in both screens (Figure 1B and Table S2). Among these ISGs were several broadly acting antiviral factors involved in pattern recognition and IFN signaling, such as *DDX58* (RIG-I), *IFIH1* (MDA5), *IRF2*, *IRF7*, *IL28RA* (IFNL1), inflammatory cytokines, *CCL5* (RANTES), and broadly acting or IAV-specific inhibitors, such as *IFITM3*, *Mx1*, and *BST2* (Schneider et al., 2014). *IFITM3* and *Mx1* act early (IAV entry or replication), whereas *BST2*, also known as tetherin, prevents release of budding virus particles at the host cell surface. Although conflicting data exist on IAV inhibition by tetherin (Mangeat et al., 2012; Watanabe et al., 2011; Yondola et al., 2011), it potently inhibited spread in A549 cells and was subsequently used as a positive control. We also identified a number of ISGs with previously uncharacterized antiviral activities: *MAP3K14*, *ELF1*, *PMM2*, *FAM46C*, *TBX3*, *SCO2*, *CRY1*, *TNFSF10* (TRAIL), *XAF1*, *MAFF*, and *SERPINE1* (serine protease inhibitor, member E1). We validated this set of genes with independently generated, high-titer lentiviral stocks and A549 cells, as well as normal human bronchial epithelial cells (NHBE). All but *FAM46C*, which did not inhibit virus spread on NHBE cells, were confirmed as antiviral (Figure 1C). To exclude false positives due to cytotoxicity, we tested cell proliferation, apoptosis, DNA damage, and cell death profiles by high-throughput microscopy (HTM) (Figure S1D). We found that only expression of *MAP3K14*, the most potent hit in our screen, and *TNFSF10* were cytotoxic relative to the empty vector control.

Because protease inhibitors have been used clinically to treat other viruses (e.g., HIV), an endogenous effector with a similar function was a promising lead. We therefore focused on exploring the antiviral action of *SERPINE1*, encoding PAI-1. *SERPINE1* expression inhibited spread of various clinical IAV isolates, including a derivative of a highly pathogenic avian H5 influenza virus, modified to remove the polybasic cleavage site in the viral hemagglutinin (Steel et al., 2009), A/Vietnam/1203/2004(HALo) (H5N1), the pandemic A/California/04/2009 (H1N1), and an isolate of swine origin, A/sw/Texas/4199-2/1998 (H3N2) (Figure 1D). In multi-step viral growth kinetics, *SERPINE1* expression reduced extracellular IAV WSN/33 titers ~10-fold, comparable to inhibition by tetherin (Figure 1E).

This versatile SERPIN family member has been implicated in many physiological processes, including regulation of fibrinolysis (reviewed in Declerck and Gils, 2013). However, since an antiviral effector function of PAI-1 protein in the context of the intrinsic immune response is novel, we set out to determine its role in restricting IAV infection.

### IAV Infection Enhances Secretion of PAI-1, which Is Both Necessary and Sufficient for IAV Inhibition

We first studied the kinetics of *SERPINE1* gene expression, as well as PAI-1 protein production and secretion. We compared

A549 cells and the more relevant in vitro model of NHBE-derived, differentiated human ciliated airway epithelium cultures (HAEc), which mimic both the morphology and physiology of the airway epithelium in vivo. In A549 cells, *SERPINE1* mRNA was slightly upregulated upon IFN- $\beta$  stimulation and following infection with IAV WSN/33 (Figure 2A). This was not due to nonresponsiveness of A549 cells, since other ISGs were highly upregulated (Figures S2A–S2C). TGF- $\beta$  is known to trigger *SERPINE1* expression via the canonical Wnt/ $\beta$ -catenin pathway (He et al., 2010). Indeed, TGF- $\beta$  treatment of A549 cells strongly induced *SERPINE1* expression with no or modest effects on *ISG15*, *IFITM3*, and *BST2* mRNA levels (Figures S2A–S2D). Stimulation of *SERPINE1* gene expression led to increased intracellular and extracellular levels of PAI-1 (Figures 2B, 2C, S2E, and S2F). Consistent with PAI-1 being efficiently secreted, total PAI-1 levels in the supernatant were about 16-fold higher than in respective IFN- $\beta$ -treated cell lysates at 24 hr (Figures 2B and 2C). We observed apical secretion of PAI-1 by HAEc after either IAV WSN/33 infection (Figure 2D) or TGF- $\beta$  treatment (Figure S2G). Of note, even mock-treated A549 cells and HAEc constantly produced and secreted basal levels of PAI-1 that accumulated over time (Figures 2B–2D). However, PAI-1 is further upregulated by certain stimuli, including virus infection.

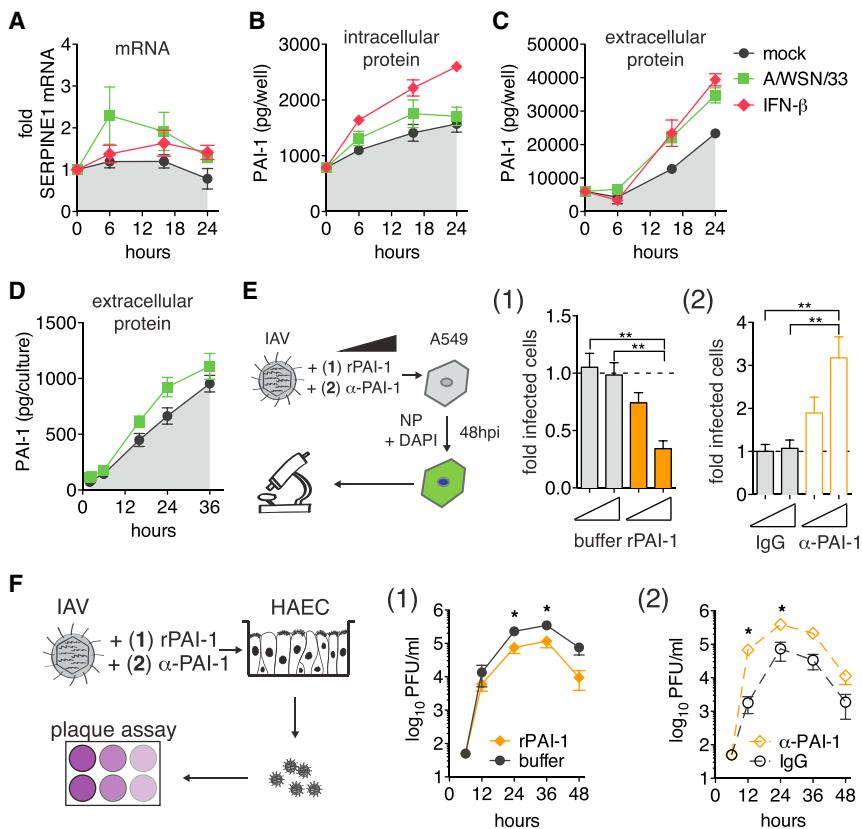
To test possible IAV inhibition by extracellular PAI-1, we added recombinant active PAI-1 (rPAI-1) to the supernatant of A549 cells during IAV infection. rPAI-1 decreased IAV WSN/33 spread in a dose-dependent manner (Figure 2E). Conversely, we used a polyclonal  $\alpha$ -PAI-1 antibody that targets the  $\alpha$ -helix F of PAI-1 to neutralize PAI-1's ability to inhibit proteases (Komissarov et al., 2005). Strikingly, we found that addition of  $\alpha$ -PAI-1 antibody to A549 cells during infection significantly enhanced IAV spread compared to the IgG control (Figure 2E).

Thus far, the extracellular environment in our experiments was defined by components of the growth medium or by proteins secreted by the cultured cells (endogenous or overexpressed). To test the potency of PAI-1 in a more natural setting, we examined IAV growth kinetics on HAEc in the presence or absence of rPAI-1. Strikingly, addition of rPAI to the apical side of HAEc significantly reduced IAV growth compared to carrier control, with ~10-fold lower infectivity at 48 hpi (Figure 2F). In contrast, addition of  $\alpha$ -PAI-1 antibody dramatically enhanced IAV growth as early as 12 hpi and continued throughout the course of infection. In both A549 cells and HAEc, human parainfluenzavirus 3 (HPIV3) was unaffected by either of these treatments, showing that the effect was selective for IAV and not due to cytotoxicity (Figures S2H and S2I).

Experiments probing which stage of the viral life cycle was affected by PAI-1 showed no effects on early stages, replication, or egress (Figures S3A–S3E). However, a reduction of progeny particle infectivity (Figure S3F), together with the extracellular presence of PAI-1 (Figure 2) and its well-described function as protease inhibitor (Declerck and Gils, 2013), led us to investigate this activity as the mechanistic basis of PAI-1-mediated IAV inhibition.

### PAI-1 Targets Airway Proteases Needed for Extracellular IAV Maturation

Infectivity of IAV progeny particles requires a maturation cleavage of the viral hemagglutinin (HA0 to HA1 and HA2), catalyzed



**Figure 2. SERPINE1 Gene Expression Profiles and the Role of Extracellular PAI-1 Protein for IAV Inhibition**

(A–C) A549 cells were infected with IAV WSN/33, treated with IFN- $\beta$ , or mock-treated. (A) SERPINE1 mRNA expression was normalized relative to housekeeping gene RPS-11. Fold increase over pre-treatment control levels is shown. Accumulated total (B) intracellular and (C) extracellular protein levels of PAI-1 were determined by ELISA. Data are represented as mean  $\pm$  SEM from  $n = 3$  experiments.

(D) HAEC infected with IAV WSN/33 or mock treated. Accumulated PAI-1 protein in repeated apical washes was determined by ELISA. Data are shown as mean  $\pm$  SEM from  $n = 3$  replicates.

(E) A549 cells were infected with IAV WSN/33 in the presence of either rPAI-1 (1) or  $\alpha$ -PAI-1 antibody (2), and virus spread was assayed by HTM at 48 hpi. The number of infected cells was normalized to buffer (1) or IgG (2). Data are represented as mean  $\pm$  SEM from  $n = 4$  independent experiments. Statistical significance relative to empty control was determined by t tests.

(F) HAECs were infected with IAV WSN/33 in the presence of either rPAI (1) or  $\alpha$ -PAI-1 (2) and buffer (1) or IgG (2). Progeny virus was collected from apical washes, and rPAI-1 or  $\alpha$ -PAI-1 were replenished after each wash. Virus titers were determined by plaque assay. Data are shown as mean  $\pm$  SEM from  $n = 3$  replicates.

See also Figure S2.

by host proteases (Lazarowitz and Choppin, 1975; Skehel and Waterfield, 1975). We tested whether known HA-cleaving proteases might be direct targets for PAI-1 inhibition. PAI-1 inhibits protease function by forming a covalent, SDS-stable bond with the target protease, detectable in gel shift assays. Urokinase plasminogen activator (uPA), a major target of PAI-1, formed a complex of  $\sim$ 85 kDa with rPAI-1, but not in the presence of the PAI-1 inhibitor triplexin (Figure 3A). TPCK-trypsin is another established target of PAI-1 and a prototype chymotrypsin-like serine protease (Olson et al., 2001). To allow multiple rounds of IAV infection in cells lacking endogenous HA-cleaving proteases, such as A549, trypsin is typically added in low concentrations to the tissue culture medium. We readily detected PAI-1-trypsin complexes formed in vitro and their degradation products, which were again absent when triplexin was added (Figure 3B). Furthermore, we identified human tryptase and human airway trypsin (HAT) as PAI-1 targets (Figures 3C and 3D). In contrast, we were unable to detect complexes of PAI-1 and furin, although this interaction has been previously reported (Figure S4) (Bernot et al., 2011).

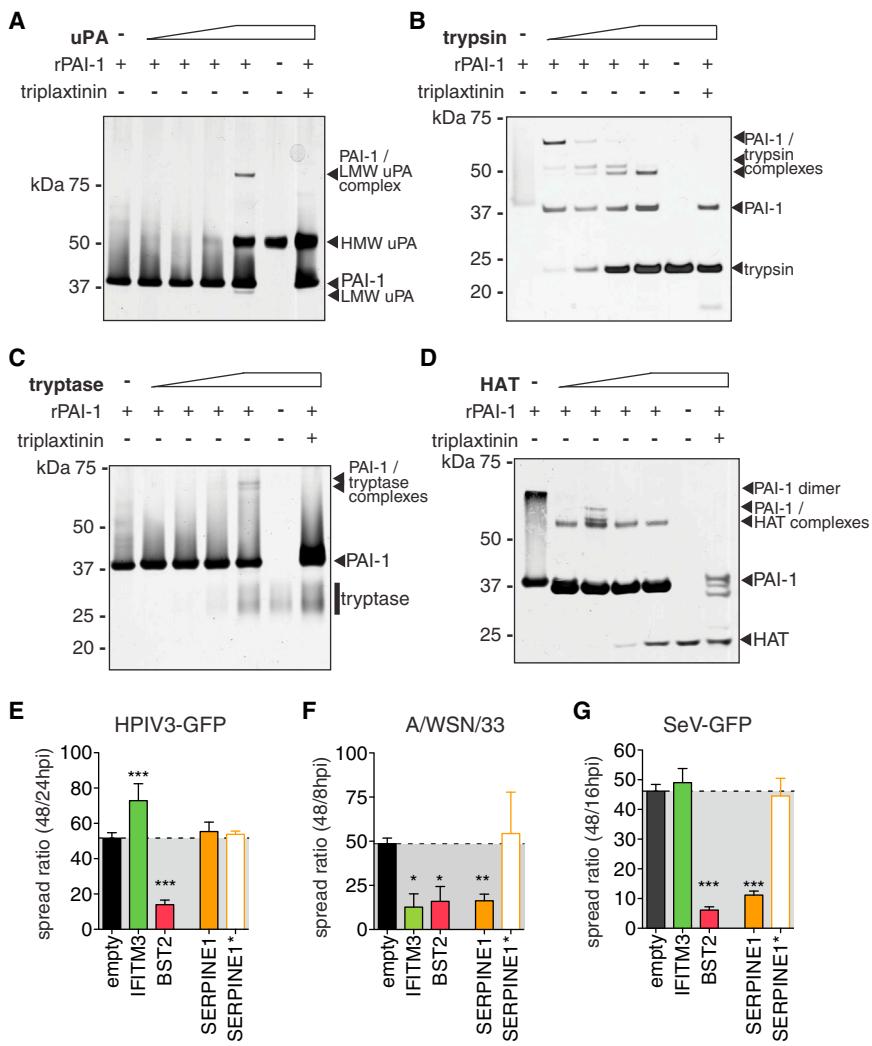
The use of host proteases for maturation is not unique to IAV. Paramyxoviruses also use airway proteases to cleave their surface fusion (F) glycoprotein, and Sendai virus (SeV) utilizes the same extracellular airway proteases as IAV (Tashiro et al., 1992). In contrast, HPIV3 F is cleaved intracellularly by ubiquitous endoproteases such as furin. Using the spread assay and trypsin as protease, we found that HPIV3 was not inhibited by PAI-1, whereas IAV WSN/33 and SeV both exhibited reduced

spread (Figures 3E–3G). SERPINE1\* encodes a catalytically inactive PAI-1 due to a single point mutation in the reactive center loop and served as a loss-of-function negative control (Lawrence et al., 1994). These data clearly demonstrate that PAI-1 targets trypsin-like airway proteases needed for extracellular virus maturation.

#### PAI-1 Prevents HA Cleavage of IAV Progeny Particles

Next, we examined the efficiency of HA cleavage in the presence of PAI-1 and trypsin. We found that rPAI-1 completely blocked trypsin-mediated cleavage of HA0 into HA1 and HA2 (Figure 4A). Furthermore, expression of wild-type PAI-1, but not the inactive mutant, reduced the specific infectivity of IAV Puerto Rico/8/34 grown in the presence of trypsin (Figure 4B). We found similar results for the extracellular airway proteases transmembrane protease serine 2 (TMPRSS2) and HAT, both known HA-cleaving proteases (Hatesuer et al., 2013; Figure 4C). Treatment with additional trypsin prior to virus titration restored specific infectivity to the level of virus grown in the presence of either protease, suggesting that the defect in specific infectivity was indeed caused by uncleaved HA (Figures 4B and 4C).

We next investigated cleavage efficiency of HA from different clades: H1 of A/Puerto Rico/8/34, H2 of A/Japan/305/57, and wild-type H5 of highly pathogenic A/Vietnam/1203/2004 (Figures 4D and 4E). Both H1<sub>PR8</sub> and H2<sub>Japan</sub> were uncleaved at baseline but were readily cleaved by either exogenous trypsin or overexpressed TMPRSS2 (Figure 4D).

**Figure 3. Protease Targets of PAI-1**

(A–D) Representative gel shift assays to show complexes of recombinant PAI-1 (rPAI-1) and indicated proteases. 1 µg of rPAI-1 was combined with increasing amounts of protease, and the mixture was separated on SDS gels followed by silver staining. Where indicated, triplaxtinin was used to inhibit PAI-1 activity.

(E–G) HTM spread assay in the presence of PAI-1 wild-type (*SERPINE1*), loss-of-function mutant (*SERPINE1\**), and controls for HPIV3-GFP (E), IAV WSN/33 (F), or SeV-GFP (G). Data are shown as mean ± SEM from at least n = 6 replicates from 2 independent experiments. One-way ANOVA and Kruskal-Wallis post-test.

See also Figure S4.

Taken together, these data show that PAI-1 inhibits IAV spread by reducing extracellular cleavage maturation of progeny particles and that this mechanism operates in a physiologically relevant model of airway epithelium.

### *Serpine1*<sup>-/-</sup> Mice Exhibit Enhanced IAV Infection and More Severe Disease Pathology

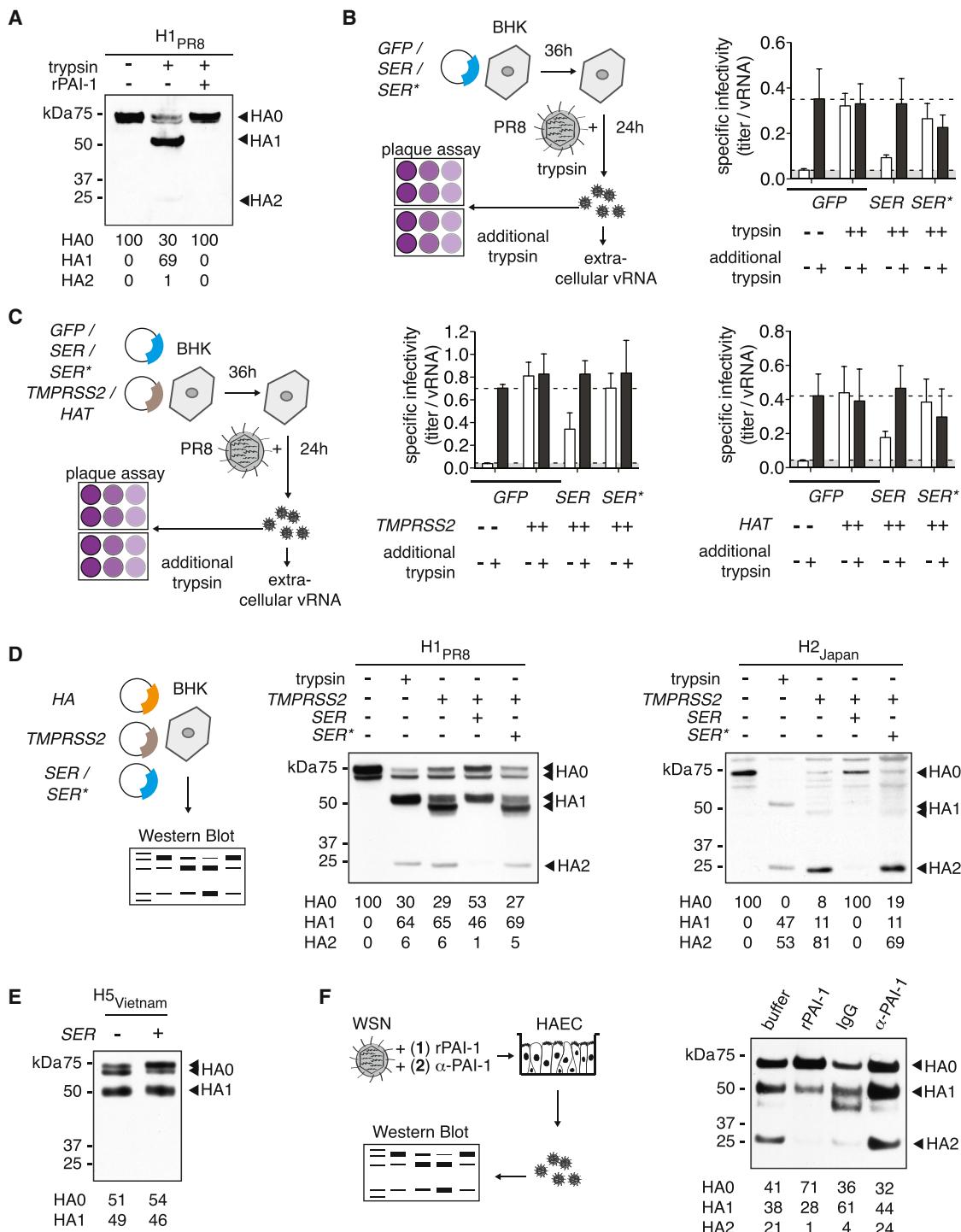
To examine the role of PAI-1 during IAV infection *in vivo*, we infected *Serpine1*<sup>-/-</sup> or *Serpine1*<sup>+/+</sup> (B6) control mice intranasally with IAVs Puerto Rico/8/34, WSN/33, or vesicular stomatitis virus (VSV) as a control. *Serpine1* was transcriptionally upregulated in B6 mouse lungs upon IAV infection (Figure 5B). We also detected increasing levels of murine PAI-1 (mPAI-1) in mouse lungs over the course

The slight differences in HA1 size between trypsin- and TMPRSS2-treated HA has been previously shown to be caused by altered glycosylation patterns in TMPRSS2-overexpressing cells (Bertram et al., 2010a). Further, we found that wild-type but not mutant PAI-1 inhibited TMPRSS2-mediated cleavage of both HA subtypes. In contrast, cleavage of H5<sub>Vietnam</sub>, which is mediated by endogenous intracellular proteases such as furin, was not impaired by PAI-1 (Figure 4E). This result is in line with data from PAI-1-furin gel shift assays (Figure S4) and from experiments with furin-dependent HPIV3 (Figures S2H and S2I), further suggesting that furin is indeed not inhibited by PAI-1.

Finally, we examined the cleavage state of HA on HAEC. These cultures produce and secrete intrinsic airway proteases on their apical side; in fact, they have previously enabled identification of several proteases that cleave HA (Böttcher et al., 2006). Addition of rPAI-1 to HAEC during IAV WSN/33 infection dramatically reduced HA cleavage, whereas addition of α-PAI-1 increased cleavage (Figure 4F). These findings are in agreement with our previous results demonstrating virus growth kinetics in HAEC (Figure 2F).

of IAV infection with a 10-fold increase at 5 dpi compared to PBS-treated B6 (Figure 5C). Both *Serpine1* mRNA and mPAI-1 protein were undetectable in *Serpine1*<sup>-/-</sup> lung homogenates. IAV-infected *Serpine1*<sup>-/-</sup> mice exhibited significant weight loss and succumbed to infection on average 1 day earlier than B6 parental mice (Figures 5D, 5E, 5G, and 5H). This difference was not observed in mice infected with VSV control (Figure S5A). Despite the modest difference in survival, lung IAV titers at late times of infection were significantly elevated in *Serpine1*<sup>-/-</sup> (Figures 5F and 5I), which correlated with the levels of mPAI-1. Keller et al. (2006) used *Serpine1*<sup>-/-</sup> mice to study the effect of influenza virus infection on thrombosis, and they found no impact of PAI-1 on IAV infection, as measured by the amount of viral RNA at day 4 dpi. However, given the role of PAI-1 in blocking viral maturation, the levels of infectious virus particles (Figures 5F and 5I) are a more relevant measure. *Serpine1*<sup>+/+</sup> mice exhibited an intermediate phenotype with respect to weight loss, survival, and murine lung PAI-1 levels, indicating a co-dominant effect of PAI-1 on IAV infection (Figure S5B).

*Serpine1*<sup>-/-</sup> lungs were increased in size at 5 days post-IAV infection, with enhanced necrosis compared to lungs from B6

**Figure 4. Effect of PAI-1 on Maturation of IAV HA**

(A, D, and E) BHK cells were transfected in serum-free MEM to express HA of different origins. Whole-cell lysates were analyzed by western blot. HA cleavage products HA0 and HA1 are indicated, and relative band intensities are shown at the bottom of the gel.

(A) HA of A/Puerto Rico/8/34 origin. Shown are cells treated with TPCK-trypsin, TPCK-trypsin and rPAI-1, or buffer.

(B) BHK cells were transfected to express GFP as negative control, PAI-1 wild-type (SER1), or loss-of-function mutant (SER\*) and then challenged with IAV Puerto Rico/8/1934 in the presence of TPCK-trypsin. Supernatants were harvested and assayed for infectivity by focus forming assay on MDCK cells and for viral RNA by qRT-PCR. Plaque assays were performed in duplicate with or without additional post-harvest incubation with TPCK-trypsin. Specific infectivity was determined by calculating the ratio of FFU to vRNA copies, and values were normalized to GFP control. Data are shown as mean ± SEM from n = 6 replicates from 2 independent experiments.

(legend continued on next page)

controls (Figure 5A). Bleeding into alveoli was slightly enhanced in infected *Serpine1*<sup>-/-</sup> lungs at 4 dpi, which might be expected given PAI-1's role in fibrinolysis (Figure S5C). Cytokine expression was elevated in IAV-infected *Serpine1*<sup>-/-</sup> lungs compared to B6 control lungs, indicating increased inflammation (Figure S5D). In contrast, cytokine levels in uninfected lungs or in lungs from mice intranasally challenged with poly(I:C) were comparable between genotypes (Figure S5E). Thus, the increased weight loss and death observed in *Serpine1*<sup>-/-</sup> mice during IAV infection cannot be simply attributed to a global upregulation of cytokines. Whereas cytokine levels began to decrease in IAV-infected B6 lungs at 5 dpi, suggesting a subsiding infection, they stagnated in *Serpine1*<sup>-/-</sup> lungs, indicating an ongoing infection in *Serpine1*<sup>-/-</sup> mice. This correlated with the presence of higher IAV titers in *Serpine1*<sup>-/-</sup> mice at this late time point.

To further confirm the in vivo phenotype, we generated murine tracheal epithelial cultures (MTEC) from the trachea of *Serpine1*<sup>-/-</sup> or B6 wild-type mice, similar to HAEC. After challenge with A/X-31(H3N2), we found significantly increased vRNA and infectious virus in the *Serpine1*<sup>-/-</sup> cultures at later times of infection (Figure 5J). These results strengthen the physiological relevance of *Serpine1* as regulator of IAV spread and a component of the host barrier that restricts IAV infection in vivo.

#### Natural PAI-1 Deficiency in Human Fibroblasts Is Correlated with Increased Spread of IAV

Several single nucleotide polymorphisms (SNPs), often linked to bleeding disorders, have been described for *SERPINE1* (Fay et al., 1997; Zhang et al., 2005). SNP rs6092 is located in the signal peptide sequence of *SERPINE1* (c.A43T), resulting in substitution of alanine by threonine (PAI-1 p.A15T) and partial intracellular retention of PAI-1. In fact, a heterozygous rs6092 carrier had 30% reduced serum PAI-1 and a tendency to hemorrhage (Zhang et al., 2005). We characterized IAV susceptibility in three human fibroblast lines with at least one rs6092 allele. Two of these are heterozygous for *SERPINE1* c.A43T (designated as T/A 4 and 5, Figure 6), and one is homozygous (A/A 6). We compared them to three control lines, all encoding wild-type PAI-1 (T/T 1-3).

We found that IAV WSN/33 spread was significantly enhanced in the three c.A43T fibroblast lines compared to controls (Figures 6A and 6B). This was not due to enhanced replication, at least for T/A 5 and A/A 6 because the number of infected cells during one round of replication was similar to controls (Figure 6C). In multicycle growth experiments, we observed some variation between control lines, which is common, as donors are genetically heterogeneous. However, we found that each c.A43T fibroblast line produced significantly more infectious virus compared to each of the three controls (Figure 6D).

To determine whether this phenotype was selective for IAV, we challenged the fibroblasts with HPIV3. Indeed, two of our three c.A43T fibroblast lines, T/A 5 and A/A 6, did not support increased spread or replication of HPIV3, as determined by HTM, or in multicycle growth assays (Figures 6E–6G). However T/A 4 exhibited both enhanced HPIV3 spread and replication compared to controls, in which we had previously observed enhanced IAV replication (Figure 6C), suggesting that this cell line may have other properties unrelated to *SERPINE1* that promote virus replication and spread.

Extracellular PAI-1 levels of c.A43T fibroblast cultures were about 50% that of controls, which was consistent with previous reports (Figure 6H; Zhang et al., 2005). We next attempted to rescue IAV spread inhibition by adding rPAI-1 to the culture medium of A/A 1, T/A 5, and A/A 6 cells during IAV WSN/33 infection. rPAI-1 dramatically reduced virus spread in T/A 5 and A/A 6 cells down to the level of T/T 1 control cells in a dose-dependent manner. In contrast, control *STAT1*<sup>-/-</sup> fibroblasts, which support enhanced spread, but do not harbor c.A43T nor have reduced extracellular PAI-1 levels, could not be rescued to the level of T/T 1 controls (Figure 6I).

These results show that a natural extracellular PAI-1 deficiency correlates with increased IAV spread.

#### DISCUSSION

We identified the serine protease inhibitor PAI-1 as an ISG that restricted IAV spread, and we characterized its previously undescribed antiviral function. PAI-1 is a 50 kDa glycoprotein and the main physiological inhibitor of urokinase/tissue plasminogen activators (uPA/tPA), both major regulators of the fibrinolytic system. uPA/tPA convert the zymogen plasminogen into plasmin, triggering a proteolytic cascade to dissolve blood clots (reviewed in Declerck and Gils, 2013). PAI-1 is present in plasma, platelets, and the extracellular matrix and is secreted by endothelial, smooth muscle, and immune cells in different tissues, including the airway. Other than uPA/tPA, PAI-1 inhibits multiple serine proteases of the chymotrypsin type, with varying efficiencies (Irving et al., 2000).

Here, we find three new PAI-1 protease targets: human tryptase (tryptase Clara; club cell secretory protein), HAT, and TMPRSS2, all of which are involved in extracellular IAV glycoprotein cleavage. This finding suggests a role for PAI-1 as an antiviral factor by targeting extracellular maturation of IAV particles.

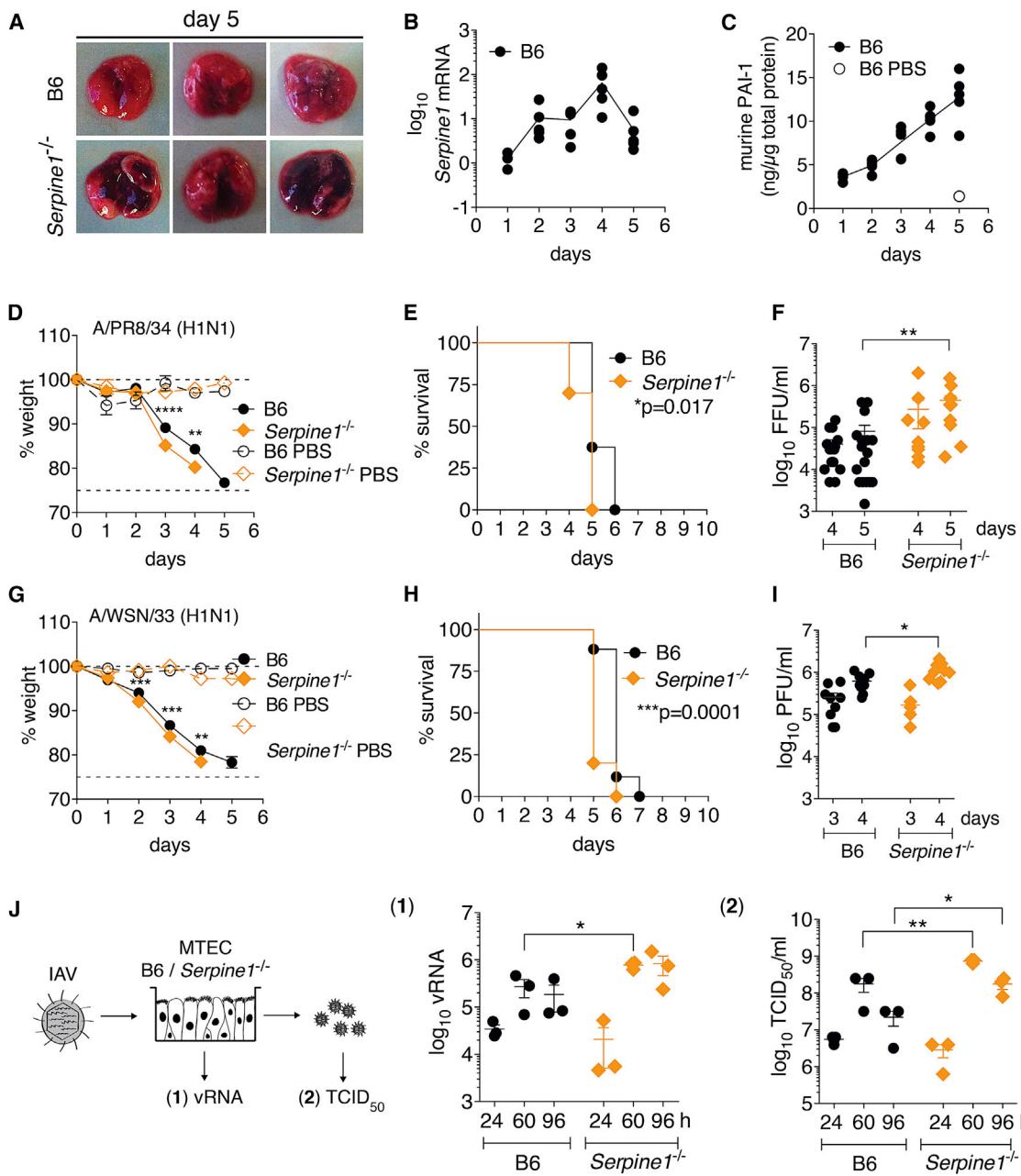
IAV maturation involves cleavage of its surface glycoprotein HA into HA1 and HA2. These proteins remain linked via a disulfide bond (Skehel and Waterfield, 1975), which forms a hinge, a prerequisite for viral envelope fusion during entry (Klenk et al., 1975; Lazarowitz and Choppin, 1975). Because IAV does not encode its own protease, HA cleavage depends on the

(C) BHK cells were transfected to co-express GFP as negative control, PAI-1 wild-type (SER), or loss-of-function mutant (SER<sup>\*</sup>), and TMPRSS2 or HAT as bait protease. Cells were then challenged, and virus infectivity was analyzed as described in (B).

(D) HA cleavage assay with H1 of A/Puerto Rico/8/34 or H2 of A/Japan/305/57 origin in the presence of protease TMPRSS2 and wild-type (SER) or loss-of-function (SER<sup>\*</sup>) PAI-1.

(E) HA cleavage assay with H5 of A/Vietnam/1203/2004 origin and wild-type (SER) PAI-1, or GFP as a negative control.

(F) HA cleavage assay on HAEC infected with IAV WSN/33 in the presence of either rPAI (1) or  $\alpha$ -PAI-1 (2) and buffer (1) or IgG (2). Progeny virus particles from apical washes at 24 hpi were analyzed.



**Figure 5. Role of Serpine1 in Lethal IAV Infection in Mice**

(A–F) Wild-type (B6) or Serpine1<sup>-/-</sup> mice were infected intranasally with 36 LD<sub>50</sub> of IAV Puerto Rico/8/34. Statistical significance was determined by individual t tests (D, F, G, and I) or log rank test (E and H).

(A) Representative images of infected lungs 5 dpi.

(B and C) Homogenates of infected mouse lungs were assayed for Serpine1 mRNA levels by qRT-PCR (B) and mPAI-1 protein levels by ELISA (C). Data are shown as mean  $\pm$  SEM from n = 5 mice per group.

(D and E) Weight loss (D) and survival (E) of infected mice. Data are shown as mean  $\pm$  SEM from n = 25 infected B6 or Serpine1<sup>-/-</sup> mice, and n = 3 respective PBS control animals.

(F) Homogenates of infected lungs were assayed for IAV titers by focus forming assay on MDCK cells. Data are shown as mean  $\pm$  SEM from n = 17 B6 and 11 Serpine1<sup>-/-</sup> mice per day.

(G–I) Wild-type (B6) or Serpine1<sup>-/-</sup> mice were infected intranasally with 36 LD<sub>50</sub> of IAV WSN/33 and monitored for weight loss (G) and survival (H). Data are shown as mean  $\pm$  SEM from n = 32 infected B6 or Serpine1<sup>-/-</sup> mice, and n = 3 respective PBS control animals. IAV lung titers (I) shown as mean  $\pm$  SEM from n = 9 B6 or Serpine1<sup>-/-</sup> mice per day.

(legend continued on next page)

presence of host proteases at the site of IAV replication. Furin and PC5/6, members of the subtilisin-like family of serine-proteases, are known to cleave HA intracellularly (Horimoto et al., 1994; Stieneke-Gröber et al., 1992). We could not find evidence for inhibition of furin by PAI-1. Human tryptase, HAT, TMPRSS2, TMPRSS13, mosaic serine protease large-form (MSPL), matrix metalloproteinase 9 (MMP9), and plasmin are known extracellular HA-cleaving proteases (Lazarowitz et al., 1973; reviewed in Bertram et al., 2010b). Most are members of the chymotrypsin-like serine protease family, the “target family” of PAI-1. Here, we provide direct proof that PAI-1 efficiently inhibits trypsin- and TMPRSS2-mediated cleavage of HA, and we hypothesize that PAI-1 might inhibit other members of this protease family as well.

Accessibility to glycoprotein cleavage is a major determinant of IAV pathogenicity (Bosch et al., 1981). Whether, and at what efficiency, an HA subtype can be cleaved by a given protease is determined by its cleavage site sequence. Highly pathogenic avian H5 and H7 HAs contain a multi-basic cleavage site, which makes them accessible to ubiquitously expressed, intracellular proteases like furin (Bosch et al., 1981). Hence, viral particles of these subtypes are immediately mature and infectious upon budding. Less pathogenic IAV HAs, in contrast, contain a single arginine residue in their cleavage site, which makes them dependent on cleavage by extracellular airway proteases, such as TMPRSS2. Variation in the amino acid sequence surrounding this arginine alters cleavage efficiency (Galloway et al., 2013). Recent studies in *TMPRSS2*<sup>-/-</sup> mice revealed a direct influence on pathogenicity of IAV strains relying on this protease (Hatesuer et al., 2013; Sakai et al., 2014; Tarnow et al., 2014). In the absence of TMPRSS2, IAV of H1N1, H3N2, or H7N9 origin were unable to spread from the trachea to the lungs, making the mice highly resistant. This was not the case for highly pathogenic H5N1 virus (Sakai et al., 2014). Thus, sensitivity to PAI-1-mediated inhibition (or inhibition by other protease inhibitors) will depend on the nature of the cleaving protease, as well as on the cleavage efficiency of a specific HA subtype by that protease (Figure 5). In this regard, it will be interesting to determine whether the protease inhibitor PAI-1 has differential impact on pathogenicity of different IAV subtypes *in vivo*.

Not surprisingly, HA cleavage has been proposed as a target for antiviral therapy. Synthetic inhibitors, such as morpholinos or peptide mimetic protease inhibitors, have been used to target single proteases (Böttcher-Friebertshäuser et al., 2011). Leupeptin of actinomycetes origin, the synthetic drug camostat, or aprotinin from bovine lungs all have broader antiprotease activity (Beppu et al., 1997; Lee et al., 1996; Tashiro et al., 1987; Zhirnov et al., 2011). So does recombinant mucus protease inhibitor (MPI), the only previously known airway protease inhibitor of human origin with anti-IAV activity (Beppu et al., 1997). The efficacy of applying these molecules to restrict IAV has been demonstrated both *in vitro* and *in vivo* and validates HA cleavage inhibition as an attractive antiviral strategy. Our data show that PAI-1

employs this mechanism and is the first known host protease inhibitor that functions to protect the host during natural IAV infection.

PAI-1 is a somewhat unconventional ISG, as it is constitutively expressed, but can be further upregulated by IFN and other cytokines like IL-6, IL-1, TGF-β, and TNF-α (Medcalf, 2007). This is significant because HA-cleaving proteases are often upregulated by IAV infection, skewing the protease-protease inhibitor balance in favor of the virus (Kido et al., 2012). Thus, PAI-1 may tip the balance back toward the host. Its regulation by TGF-β might be particularly relevant during IAV infections because IAV neuraminidase can activate latent TGF-β (Carlson et al., 2010). Other sources of TGF-β in the airway during infection include secretion by macrophages and by epithelial cells.

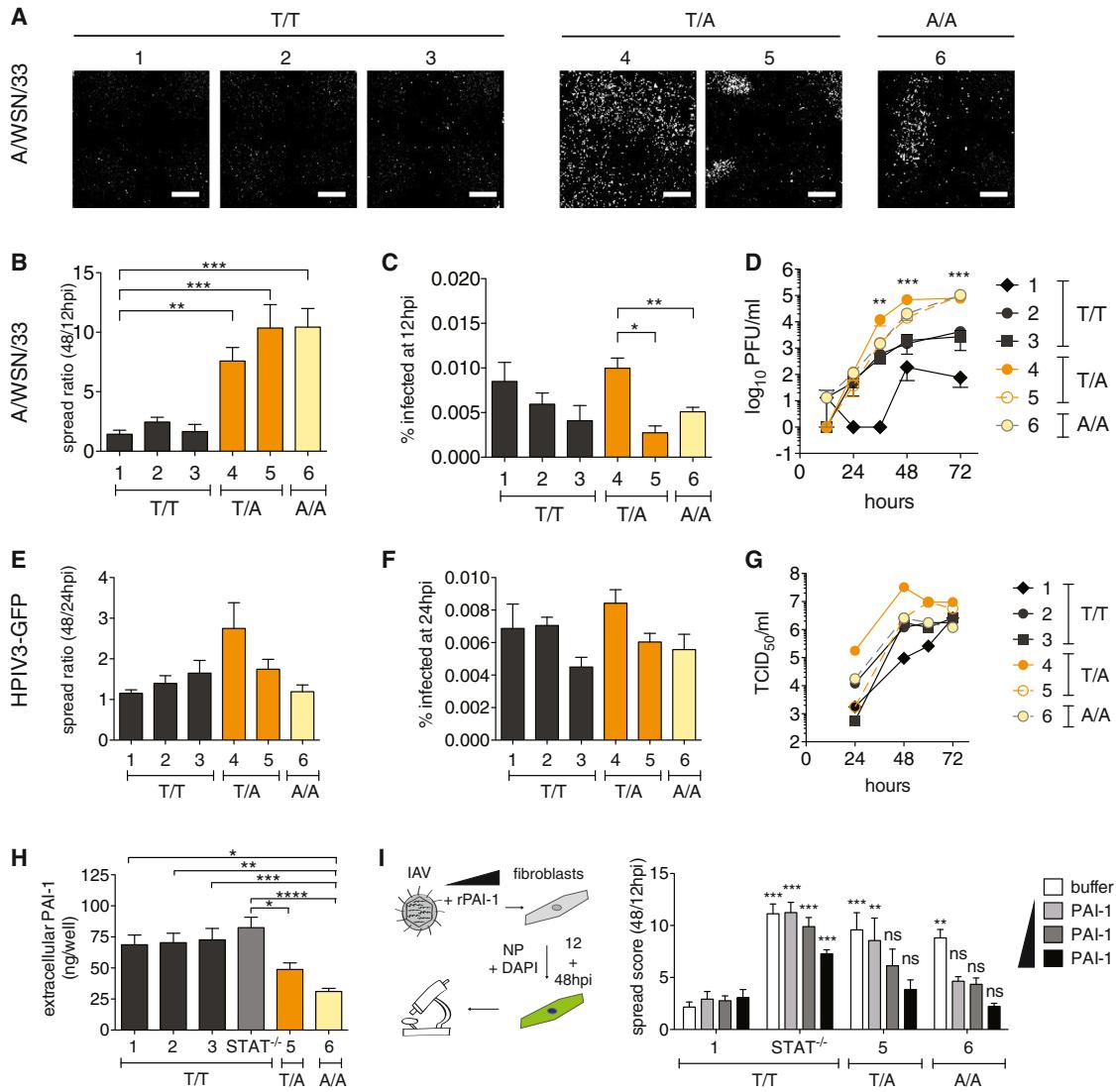
It is interesting that upregulation of *SERPINE1* gene expression during infection is not as dramatic as for other ISGs, possibly reflecting PAI-1's critical role in other local and systemic physiological processes, where massive overproduction could be deleterious for blood clotting, cancer metastasis, or cell adhesion and migration (reviewed in Declerck and Gils, 2013). PAI-1 is the only SERPIN that spontaneously—and quickly—adopts a latent form *in vivo*, which provides a regulatory mechanism, both spatially and temporally (Berkenpas et al., 1995).

The respiratory tract is a major portal for virus entry into the body but is protected by a multilayered antiviral fence. Mucus acts as a physical extracellular barrier. Mucins and surfactant proteins in mucus trap and aggregate virus particles outside of cells or act as decoy receptors inhibiting early steps in the viral replication cycle (Hartshorn et al., 2006; Reading et al., 2008). Mucins are present in airways at constitutive levels but are upregulated during inflammation (Turner and Jones, 2009). We now add PAI-1 as an additional component of this extracellular barrier. PAI-1 is the first extracellular directly antiviral ISG, and its promiscuity suggests that it may also play an important role in host defense against other respiratory viruses that rely on this step in their life cycles.

Many virus families require a maturation cleavage of viral surface glycoproteins, generally realized by serine proteases. Some use virus-encoded proteases, such as picornaviruses or retroviruses (Brody et al., 1992; Lee et al., 1993). Others, such as paramyxoviruses, orthomyxoviruses, coronaviruses, filoviruses, or arenaviruses, rely on host proteases (Glowacka et al., 2011; Lenz et al., 2001; Skehel and Waterfield, 1975; Steinhauer and Plemper, 2012; Volchkov et al., 1998). All are potentially sensitive to inhibition of their activating proteases by host or pharmacological protease inhibitors. Indeed, we show that a single inhibitor, PAI-1, significantly inhibits spread of IAV and SeV. Further exploration of the virus—activating protease—host inhibitor troika may open new avenues for antiviral intervention and deepen our understanding of virus maturation and pathogenicity.

(J) MTECs cultured from B6 or *Serpine1*<sup>-/-</sup> mice were infected with A/X-31(H3N2) at MOI = 10<sup>-6</sup>. vRNA was quantified from lysed MTEC by qRT-PCR. In parallel, infectious virus titers were determined by TCID<sub>50</sub> assay on MDCK cells. Data are shown as mean ± SEM from n = 3 B6 or *Serpine1*<sup>-/-</sup> MTECs. Statistical significance was determined by two-way ANOVA.

See also Figure S5.

**Figure 6. Susceptibility of PAI-1-Deficient Human Fibroblasts to IAV**

(A–H) SV-40-immortalized human fibroblasts from different donors, wild-type (T/T 1–3), heterozygous (T/A 4 and 5), or homozygous (A/A 6) for SERPINE1 SNP rs6092 were infected with IAV WSN/33 (A–D, H, and I) or HPIV3-GFP (E–G).

(A) Representative images from single wells infected with IAV WSN/33 at 48 hpi.

(B) IAV spread shown as the ratio between number of infected cells at 12/48 hpi.

(C) Percent of infected fibroblasts after one round of replication. Data in (A) and (B) are represented as mean ± SEM from n = 4 independent experiments, each in triplicate.

(D) IAV WSN/33 growth on fibroblasts. PFU in the supernatants were determined by plaque assay on MDCK cells. Data are represented as mean ± SEM from n = 4 independent experiments. Statistical significance was determined by individual t test, comparing the pooled SNP-carrying cells with the pooled control cells TT and TT3 for every time point.

(E) HPIV3-GFP spread shown as the ratio between number of infected cells at 48 and 24 hpi.

(F) Percent of infected fibroblasts after one round of replication. Data in (E) and (F) are represented as mean ± SEM from n = 2 independent experiments, each in triplicate.

(G) HPIV3-GFP growth on fibroblasts. TCID<sub>50</sub> in supernatants were determined on LLC-MK2 cells.

(H) Extracellular PAI-1 levels from IAV WSN/33-infected fibroblasts during spread assay at 48 hpi. Data are shown as mean ± SEM from total n = 8 replicates from n = 4 experiments.

(I) Rescue of virus inhibition on selected human fibroblast lines by adding rPAI-1 to the cultures during IAV WSN/33 infection. Maximum addition of rPAI-1 (black bars) was 500 ng, minimum addition (light gray bars) 125 ng. Data are represented as mean ± SEM from n = 4 independent experiments, each in quadruplicate. One-way ANOVA and Dunnett's multiple comparison post-test against T/T 1-values of the same condition.

In conclusion, we identify *SERPINE1/PAI-1* as a host factor inhibiting IAV spread and show that PAI-1 mechanistically acts by blocking maturation of progeny IAV particles, thus reducing particle infectivity. Naturally occurring human genetic variations in PAI-1 impact the ability of IAV to spread from cell to cell, thus identifying PAI-1 as a candidate human gene that may influence susceptibility to IAV and the outcome of infection. In this regard, it will be of immediate interest to determine whether PAI-1 deficiency might be linked with the severity of human IAV infection. Finally, our work suggests that localized administration of PAI-1 to the respiratory tract might provide a new therapeutic approach for treating IAV and other respiratory viruses that require extracellular protease-driven maturation.

## EXPERIMENTAL PROCEDURES

### High-Throughput Microscopy Screening of Host Factor Library

Using Gateway technology, we transferred a previously published library of ISGs (Schoggins et al., 2011) from ORFEXPRESS entry clones into a new lentiviral expression vector (pSCRPSY), co-expressing TagRFP as a transduction control and puromycin resistance. A549 cells in 96-well plates were first transduced and then challenged with IAV WSN/33 at MOI 0.01. Due to varying transduction efficiency between the lentiviral stocks, the percentage of non-transduced cells was variable. IAV production from non-transduced cells masked potential inhibitory effects of ISGs and was eliminated by adding puromycin 1 hr prior to IAV infection. Lentiviral transduction was performed in duplicate plates for each ISG, where one sample was fixed after one round of replication (8 hpi), defining the number of initial producer cells, and the other after several rounds of replication (24 hpi). “Spread ratio” was calculated by dividing the IAV-infected cells at 24 hpi by 8 hpi for each individual ISG over empty vector-transduced cells (Figures 1A, S1A, and S1B) and yielded results that were stable over a wide range of transduction efficiencies (Figure S1B). Infected cells were detected by staining for IAV nucleoprotein (NP) and HTM (Figures 1A, S1A, and S1B).

### Mice

Mice were housed in an AAALAC-accredited facility in accordance with the Guide for the Care and Use of Laboratory Animals. All procedures outlined in the study were approved by The Rockefeller University’s Institutional Animal Care and Use Committee. Wild-type (C57BL/6J) and *Serpine<sup>-/-</sup>* (B6.129S2-Serpine<sup>1tm1Mlg/J</sup>) breeder pairs were purchased from Jackson Laboratory, and colonies were established in-house. Six- to ten-week-old mice of both sexes were used in IAV challenge, and 10- to 14-week-old mice were used for VSV experiments.

### Human Cell Lines

Informed consent was obtained from human subjects according to local regulations. Samples were subsequently transferred for experimental testing under respective ethics approvals at INSERM in France and at Rockefeller University in the USA. All protocols involving the use of human tissue were reviewed and exempted by The Rockefeller University Institutional Review Board.

### SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, five figures, and three tables and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2015.01.040>.

### AUTHOR CONTRIBUTIONS

H.-H.H. and M.A.S. contributed equally to this work. M.D. and C.M.R. designed the project. M.D. wrote the manuscript. M.D., H.-H.H., M.A.S., K.L.B., S.C., A.W., and C.M.R. designed the experiments. M.D., H.-H.H.,

M.A.S., R.H.G., K.L.B., B.F., Y.Y., S.C., and J.W.X. performed the experimental work. M.D., M.A.S., H.-H.H., S.J.C., and A.W. analyzed the results. S.J.W., M.C., A.W., P.D.B., and J.-L.C. contributed reagents and expertise.

### ACKNOWLEDGMENTS

We would like to thank all Rice lab members but especially M. Hsu, M. Saeed, and Z. Ozair, for critical reading and helpful comments on the manuscript; Jeanne Chiaravalli-Giganti at the Rockefeller High-Throughput and Spectroscopy Resource Center and Sébastien Monettes at the Tri-Institutional Laboratory of Comparative Pathology for technical advice and support; Margaret MacDonald for VSV strain Indiana; Adolfo García-Sastre for the gift of 2G9  $\alpha$ -HA antibody; Peter Palese for influenza viruses A/Puerto Rico/8/1934, A/Vietnam/1203/2004(HALO), A/California/04/2009, and A/sw/Texas/4199-2/1998; D. Steinhauer and S. Galloway for both the gift of pCAGGS-HA plasmids and for technical advice on the cleavage assay; Gary Whittaker for the gift of HAT and TMPRSS2 expression plasmids; and Sharmila Nair for advice with VSV in vivo experiments. This work was supported by National Institutes of Health (NIH) grant R01-AI091707 (C.M.R.), U54-AI057158 (I. Lipkin, Northeast Biodefense Center, subcontract to C.M.R.), the German Research Foundation and the Rockefeller Women and Science Fellowship (M.D.), Medical Research Council Grants U117597139 (to A.W. and S.C.), NIH grant 8UL1TR000043 from the National Center for Translational Sciences (NCATS; J.-L.C.), and the St. Giles Foundation (J.-L.C.). Additional funding was provided by the Greenberg Medical Research Institute, the Starr Foundation, and anonymous donors.

Received: September 19, 2014

Revised: October 22, 2014

Accepted: January 13, 2015

Published: February 12, 2015

### REFERENCES

- Beppu, Y., Imamura, Y., Tashiro, M., Towatari, T., Ariga, H., and Kido, H. (1997). Human mucus protease inhibitor in airway fluids is a potential defensive compound against infection with influenza A and Sendai viruses. *J. Biochem.* 121, 309–316.
- Berkenpas, M.B., Lawrence, D.A., and Ginsburg, D. (1995). Molecular evolution of plasminogen activator inhibitor-1 functional stability. *EMBO J.* 14, 2969–2977.
- Bernot, D., Stalin, J., Stocker, P., Bonardo, B., Scroyen, I., Alessi, M.C., and Peiretti, F. (2011). Plasminogen activator inhibitor 1 is an intracellular inhibitor of furin proprotein convertase. *J. Cell Sci.* 124, 1224–1230.
- Bertram, S., Glowacka, I., Blazejewska, P., Soilleux, E., Allen, P., Danisch, S., Steffen, I., Choi, S.Y., Park, Y., Schneider, H., et al. (2010a). TMPRSS2 and TMPRSS4 facilitate trypsin-independent spread of influenza virus in Caco-2 cells. *J. Virol.* 84, 10016–10025.
- Bertram, S., Glowacka, I., Steffen, I., Kühl, A., and Pöhlmann, S. (2010b). Novel insights into proteolytic cleavage of influenza virus hemagglutinin. *Rev. Med. Virol.* 20, 298–310.
- Bosch, F.X., Garten, W., Klenk, H.D., and Rott, R. (1981). Proteolytic cleavage of influenza virus hemagglutinins: primary structure of the connecting peptide between HA1 and HA2 determines proteolytic cleavability and pathogenicity of Avian influenza viruses. *Virology* 113, 725–735.
- Böttcher, E., Matrosovich, T., Beyerle, M., Klenk, H.D., Garten, W., and Matrosovich, M. (2006). Proteolytic activation of influenza viruses by serine proteases TMPRSS2 and HAT from human airway epithelium. *J. Virol.* 80, 9896–9898.
- Böttcher-Friebertshäuser, E., Stein, D.A., Klenk, H.D., and Garten, W. (2011). Inhibition of influenza virus infection in human airway cell cultures by an anti-sense peptide-conjugated morpholino oligomer targeting the hemagglutinin-activating protease TMPRSS2. *J. Virol.* 85, 1554–1562.
- Brody, B.A., Rhee, S.S., Sommerfelt, M.A., and Hunter, E. (1992). A viral protease-mediated cleavage of the transmembrane glycoprotein of Mason-Pfizer

monkey virus can be suppressed by mutations within the matrix protein. *Proc. Natl. Acad. Sci. USA* 89, 3443–3447.

Carlson, C.M., Turpin, E.A., Moser, L.A., O'Brien, K.B., Cline, T.D., Jones, J.C., Tumpey, T.M., Katz, J.M., Kelley, L.A., Gauldie, J., and Schultz-Cherry, S. (2010). Transforming growth factor- $\beta$ : activation by neuraminidase and role in highly pathogenic H5N1 influenza pathogenesis. *PLoS Pathog.* 6, e1001136.

Declerck, P.J., and Gils, A. (2013). Three decades of research on plasminogen activator inhibitor-1: a multifaceted serpin. *Semin. Thromb. Hemost.* 39, 356–364.

Fay, W.P., Parker, A.C., Condrey, L.R., and Shapiro, A.D. (1997). Human plasminogen activator inhibitor-1 (PAI-1) deficiency: characterization of a large kindred with a null mutation in the PAI-1 gene. *Blood* 90, 204–208.

Galloway, S.E., Reed, M.L., Russell, C.J., and Steinhauer, D.A. (2013). Influenza HA subtypes demonstrate divergent phenotypes for cleavage activation and pH of fusion: implications for host range and adaptation. *PLoS Pathog.* 9, e1003151.

Glowacka, I., Bertram, S., Müller, M.A., Allen, P., Soilleux, E., Pfefferle, S., Steffen, I., Tsegaye, T.S., He, Y., Gnrss, K., et al. (2011). Evidence that TMPRSS2 activates the severe acute respiratory syndrome coronavirus spike protein for membrane fusion and reduces viral control by the humoral immune response. *J. Virol.* 85, 4122–4134.

Hartshorn, K.L., White, M.R., Tecle, T., Holmskov, U., and Crouch, E.C. (2006). Innate defense against influenza A virus: activity of human neutrophil defensins and interactions of defensins with surfactant protein D. *J. Immunol.* 176, 6962–6972.

Hatesuer, B., Bertram, S., Mehnert, N., Bahgat, M.M., Nelson, P.S., Pöhlman, S., and Schughart, K. (2013). Tmprss2 is essential for influenza H1N1 virus pathogenesis in mice. *PLoS Pathog.* 9, e1003774.

He, W., Tan, R., Dai, C., Li, Y., Wang, D., Hao, S., Kahn, M., and Liu, Y. (2010). Plasminogen activator inhibitor-1 is a transcriptional target of the canonical pathway of Wnt/beta-catenin signaling. *J. Biol. Chem.* 285, 24665–24675.

Horimoto, T., Nakayama, K., Smeekens, S.P., and Kawaoka, Y. (1994). Protein-processing endoproteases PC6 and furin both activate hemagglutinin of virulent avian influenza viruses. *J. Virol.* 68, 6074–6078.

Irving, J.A., Pike, R.N., Lesk, A.M., and Whisstock, J.C. (2000). Phylogeny of the serpin superfamily: implications of patterns of amino acid conservation for structure and function. *Genome Res.* 10, 1845–1864.

Keller, T.T., van der Sluijs, K.F., de Kroif, M.D., Gerdes, V.E., Meijers, J.C., Florquin, S., van der Poll, T., van Gorp, E.C., Brandjes, D.P., Büller, H.R., and Levi, M. (2006). Effects on coagulation and fibrinolysis induced by influenza in mice with a reduced capacity to generate activated protein C and a deficiency in plasminogen activator inhibitor type 1. *Circ. Res.* 99, 1261–1269.

Kido, H., Okumura, Y., Takahashi, E., Pan, H.Y., Wang, S., Yao, D., Yao, M., Chida, J., and Yano, M. (2012). Role of host cellular proteases in the pathogenesis of influenza and influenza-induced multiple organ failure. *Biochim. Biophys. Acta* 1824, 186–194.

Klenk, H.D., Rott, R., Orlich, M., and Blödorn, J. (1975). Activation of influenza A viruses by trypsin treatment. *Virology* 68, 426–439.

Komissarov, A.A., Andreasen, P.A., Bødker, J.S., Declerck, P.J., Anagli, J.Y., and Shore, J.D. (2005). Additivity in effects of vitronectin and monoclonal antibodies against alpha-helix F of plasminogen activator inhibitor-1 on its reactions with target proteinases. *J. Biol. Chem.* 280, 1482–1489.

Lawrence, D.A., Olson, S.T., Palaniappan, S., and Ginsburg, D. (1994). Serpin reactive center loop mobility is required for inhibitor function but not for enzyme recognition. *J. Biol. Chem.* 269, 27657–27662.

Lazarowitz, S.G., and Choppin, P.W. (1975). Enhancement of the infectivity of influenza A and B viruses by proteolytic cleavage of the hemagglutinin polypeptide. *Virology* 68, 440–454.

Lazarowitz, S.G., Goldberg, A.R., and Choppin, P.W. (1973). Proteolytic cleavage by plasmin of the HA polypeptide of influenza virus: host cell activation of serum plasminogen. *Virology* 56, 172–180.

Lee, W.M., Monroe, S.S., and Rueckert, R.R. (1993). Role of maturation cleavage in infectivity of picornaviruses: activation of an infectosome. *J. Virol.* 67, 2110–2122.

Lee, M.G., Kim, K.H., Park, K.Y., and Kim, J.S. (1996). Evaluation of anti-influenza effects of camostat in mice infected with non-adapted human influenza viruses. *Arch. Virol.* 141, 1979–1989.

Lenz, O., ter Meulen, J., Klenk, H.D., Seidah, N.G., and Garten, W. (2001). The Lassa virus glycoprotein precursor GP-C is proteolytically processed by subtilase SKI-1/S1P. *Proc. Natl. Acad. Sci. USA* 98, 12701–12705.

Mangeat, B., Cavagliotti, L., Lehmann, M., Gers-Huber, G., Kaur, I., Thomas, Y., Kaiser, L., and Piguet, V. (2012). Influenza virus partially counteracts restriction imposed by tetherin/BST-2. *J. Biol. Chem.* 287, 22015–22029.

Medcalf, R.L. (2007). Fibrinolysis, inflammation, and regulation of the plasminogen activating system. *J. Thromb. Haemost.* 5 (1), 132–142.

Neil, S.J.D., Zang, T., and Bieniasz, P.D. (2008). Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. *Nature* 451, 425–430.

Ny, T., Sawdye, M., Lawrence, D., Millan, J.L., and Loskutoff, D.J. (1986). Cloning and sequence of a cDNA coding for the human beta-migrating endothelial-cell-type plasminogen activator inhibitor. *Proc. Natl. Acad. Sci. USA* 83, 6776–6780.

Olson, S.T., Swanson, R., Day, D., Verhamme, I., Kvassman, J., and Shore, J.D. (2001). Resolution of Michaelis complex, acylation, and conformational change steps in the reactions of the serpin, plasminogen activator inhibitor-1, with tissue plasminogen activator and trypsin. *Biochemistry* 40, 11742–11756.

Reading, P.C., Bozza, S., Gilbertson, B., Tate, M., Moretti, S., Job, E.R., Crouch, E.C., Brooks, A.G., Brown, L.E., Bottazzi, B., et al. (2008). Antiviral activity of the long chain pentraxin PTX3 against influenza viruses. *J. Immunol.* 180, 3391–3398.

Sakai, K., Ami, Y., Tahara, M., Kubota, T., Anraku, M., Abe, M., Nakajima, N., Sekizuka, T., Shirato, K., Suzuki, Y., et al. (2014). The host protease TMPRSS2 plays a major role in in vivo replication of emerging H7N9 and seasonal influenza viruses. *J. Virol.* 88, 5608–5616.

Schneider, W.M., Chevillotte, M.D., and Rice, C.M. (2014). Interferon-stimulated genes: a complex web of host defenses. *Annu. Rev. Immunol.* 32, 513–545.

Schoggins, J.W., Wilson, S.J., Panis, M., Murphy, M.Y., Jones, C.T., Bieniasz, P., and Rice, C.M. (2011). A diverse range of gene products are effectors of the type I interferon antiviral response. *Nature* 472, 481–485.

Schoggins, J.W., MacDuff, D.A., Imanaka, N., Gainey, M.D., Shrestha, B., Eitson, J.L., Mar, K.B., Richardson, R.B., Ratashny, A.V., Litvak, V., et al. (2014). Pan-viral specificity of IFN-induced genes reveals new roles for cGAS in innate immunity. *Nature* 505, 691–695.

Skehel, J.J., and Waterfield, M.D. (1975). Studies on the primary structure of the influenza virus hemagglutinin. *Proc. Natl. Acad. Sci. USA* 72, 93–97.

Steel, J., Lowen, A.C., Pena, L., Angel, M., Solórzano, A., Albrecht, R., Perez, D.R., García-Sastre, A., and Palese, P. (2009). Live attenuated influenza viruses containing NS1 truncations as vaccine candidates against H5N1 highly pathogenic avian influenza. *J. Virol.* 83, 1742–1753.

Steinhauer, D.A., and Plemper, R.K. (2012). Structure of the primed paramyxovirus fusion protein. *Proc. Natl. Acad. Sci. USA* 109, 16404–16405.

Stieneke-Gröber, A., Vey, M., Angliker, H., Shaw, E., Thomas, G., Roberts, C., Klenk, H.D., and Garten, W. (1992). Influenza virus hemagglutinin with multibasic cleavage site is activated by furin, a subtilisin-like endoprotease. *EMBO J.* 11, 2407–2414.

Tarnow, C., Engels, G., Arendt, A., Schwalm, F., Sediri, H., Preuss, A., Nelson, P.S., Garten, W., Klenk, H.D., Gabriel, G., and Böttcher-Friebertshäuser, E. (2014). TMPRSS2 is a host factor that is essential for pneumotropism and pathogenicity of H7N9 influenza A virus in mice. *J. Virol.* 88, 4744–4751.

Tashiro, M., Klenk, H.D., and Rott, R. (1987). Inhibitory effect of a protease inhibitor, leupeptin, on the development of influenza pneumonia, mediated by concomitant bacteria. *J. Gen. Virol.* 68, 2039–2041.

Tashiro, M., Yokogoshi, Y., Tobita, K., Seto, J.T., Rott, R., and Kido, H. (1992). Trypsin Clara, an activating protease for Sendai virus in rat lungs, is involved in pneumopathogenicity. *J. Virol.* 66, 7211–7216.

Turner, J., and Jones, C.E. (2009). Regulation of mucin expression in respiratory diseases. *Biochem. Soc. Trans.* 37, 877–881.

Volchkov, V.E., Feldmann, H., Volchkova, V.A., and Klenk, H.D. (1998). Processing of the Ebola virus glycoprotein by the proprotein convertase furin. *Proc. Natl. Acad. Sci. USA* 95, 5762–5767.

Wang, X., Hinson, E.R., and Cresswell, P. (2007). The interferon-inducible protein viperin inhibits influenza virus release by perturbing lipid rafts. *Cell Host Microbe* 2, 96–105.

Watanabe, R., Leser, G.P., and Lamb, R.A. (2011). Influenza virus is not restricted by tetherin whereas influenza VLP production is restricted by tetherin. *Virology* 417, 50–56.

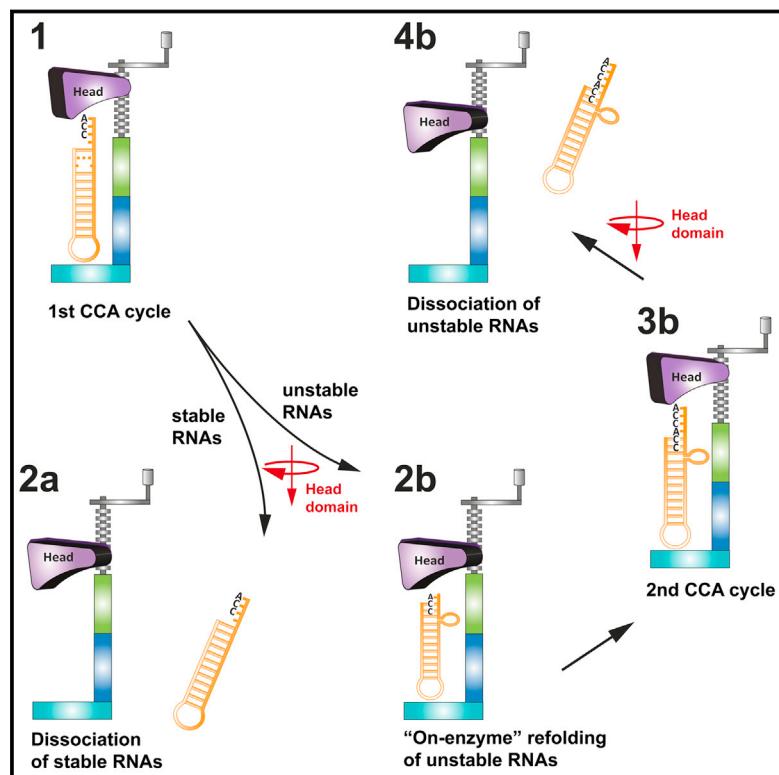
Yondola, M.A., Fernandes, F., Belicha-Villanueva, A., Uccelini, M., Gao, Q., Carter, C., and Palese, P. (2011). Budding capability of the influenza virus neuraminidase can be modulated by tetherin. *J. Virol.* 85, 2480–2491.

Zhang, Z.Y., Wang, Z.Y., Dong, N.Z., Bai, X., Zhang, W., and Ruan, C.G. (2005). A case of deficiency of plasma plasminogen activator inhibitor-1 related to Ala15Thr mutation in its signal peptide. *Blood Coagul. Fibrinolysis* 16, 79–84.

Zhirnov, O.P., Matrosovich, T.Y., Matrosovich, M.N., and Klenk, H.D. (2011). Aprotinin, a protease inhibitor, suppresses proteolytic activation of pandemic H1N1v influenza virus. *Antivir. Chem. Chemother.* 21, 169–174.

# On-Enzyme Refolding Permits Small RNA and tRNA Surveillance by the CCA-Adding Enzyme

## Graphical Abstract



## Authors

Claus-D. Kuhn, Jeremy E. Wilusz, ..., Peter A. Beal, Leemor Joshua-Tor

## Correspondence

leemor@cshl.edu

## In Brief

The CCA-adding enzyme acts as a molecular vise to produce torque on RNA substrates, and the RNAs proofread themselves through differential responses to this interrogation between stable and unstable substrates.

## Highlights

- The CCA-adding enzyme interrogates substrates like a nucleotide-driven molecular vise
- Unstable RNAs refold “on-enzyme” after completion of the first CCA cycle
- An RNA bulge permits addition of a second CCA triplet
- RNAs proofread themselves through differential responses to the CCA-adding process

## Accession Numbers

4X4N  
4X4O  
4X4P  
4X4Q  
4X4R  
4X4S  
4X4T  
4X4U  
4X4V  
4X4W

# On-Enzyme Refolding Permits Small RNA and tRNA Surveillance by the CCA-Adding Enzyme

Claus-D. Kuhn,<sup>1,4</sup> Jeremy E. Wilusz,<sup>2</sup> Yuxuan Zheng,<sup>3</sup> Peter A. Beal,<sup>3</sup> and Leemor Joshua-Tor<sup>1,\*</sup>

<sup>1</sup>W.M. Keck Structural Biology Laboratory, Howard Hughes Medical Institute, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY 11724, USA

<sup>2</sup>Department of Biochemistry and Biophysics, University of Pennsylvania Perelman School of Medicine, 415 Curie Boulevard, Philadelphia, PA 19104, USA

<sup>3</sup>Department of Chemistry, University of California, Davis, 1 Shields Ave, Davis, CA 95616, USA

<sup>4</sup>Present address: Elite Network of Bavaria, BIOMac Research Center, University of Bayreuth, 95440 Bayreuth, Germany

\*Correspondence: leemor@cshl.edu

<http://dx.doi.org/10.1016/j.cell.2015.01.005>

## SUMMARY

Transcription in eukaryotes produces a number of long noncoding RNAs (lncRNAs). Two of these, MALAT1 and Men $\beta$ , generate a tRNA-like small RNA in addition to the mature lncRNA. The stability of these tRNA-like small RNAs and bona fide tRNAs is monitored by the CCA-adding enzyme. Whereas CCA is added to stable tRNAs and tRNA-like transcripts, a second CCA repeat is added to certain unstable transcripts to initiate their degradation. Here, we characterize how these two scenarios are distinguished. Following the first CCA addition cycle, nucleotide binding to the active site triggers a clockwise screw motion, producing torque on the RNA. This ejects stable RNAs, whereas unstable RNAs are refolded while bound to the enzyme and subjected to a second CCA catalytic cycle. Intriguingly, with the CCA-adding enzyme acting as a molecular vise, the RNAs proofread themselves through differential responses to its interrogation between stable and unstable substrates.

## INTRODUCTION

The CCA-adding enzyme adds the nucleotide triplet CCA to the 3' end of all tRNAs (Deutscher, 1982), a step that is essential for tRNA aminoacylation (Sprinzl and Cramer, 1979) and correct tRNA positioning in the ribosome (Nissen et al., 2000). It is an intriguing member of the nucleotidyltransferase family, since it operates without a nucleic acid template and without translocation along the tRNA (Shi et al., 1998a). Instead, the tRNA 3'-end progressively refolds during synthesis and nucleotide selectivity switches from cytidine to adenosine after the addition of two C-nucleotides (Tomita et al., 2006; Xiong and Steitz, 2004). In addition to acting upon newly synthesized tRNA transcripts, the CCA-adding enzyme is able to accurately repair partially degraded tRNA 3' ends. Interestingly, CCA-adding enzymes fall into two divergent classes with only the catalytic domains sharing significant homology. In class I enzymes (present in

archaea), bound tRNA is involved in the selection of the correct nucleotides (Xiong and Steitz, 2004), whereas in class II enzymes (in eubacteria and eukaryotes), proper nucleotides are selected by protein only (Li et al., 2002; Tomita et al., 2004).

Although the CCA-adding enzyme was long thought to terminate polymerization once a single CCA triplet had been added, recent work on tRNA-like small RNAs derived from long non-coding RNAs (lncRNAs) indicated this is not always the case (Sunwoo et al., 2009; Wilusz et al., 2011). Metastasis-associated long adenocarcinoma transcript 1 (MALAT1) and Men $\beta$  (also known as NEAT1) are well-characterized long, nuclear-retained non-coding transcripts that are involved in cancer progression (Ji et al., 2003; Lin et al., 2007) and paraspeckle formation (Clemson et al., 2009; Imamura et al., 2014; Sunwoo et al., 2009), respectively. In addition, both of these lncRNAs contain sequences that mimic the tRNA fold and are processed by the canonical tRNA biogenesis machinery to generate small RNAs (Wilusz et al., 2008; Wilusz et al., 2011). Just as for bona fide tRNAs, the trinucleotide CCA is post-transcriptionally added to the 3'-end of MALAT1-associated small cytoplasmic RNA (mascRNA), the tRNA-like transcript originating from the 3'-end of MALAT1 (Wilusz et al., 2008). In contrast, the Men $\beta$  tRNA-like small RNA was found to be subjected to CCACCA addition and efficiently degraded (Sunwoo et al., 2009; Wilusz et al., 2011). Unlike the acceptor stems of canonical tRNAs and mascRNA, the Men $\beta$  acceptor stem is destabilized through mismatches or wobble base pairs, somehow prompting the CCA-adding enzyme to repeat its catalytic cycle, thereby adding tandem CCA motifs (Wilusz et al., 2011). The CCACCA tail then serves as a degradation signal for the cellular RNA decay machinery.

It is now clear that CCA-adding enzymes from all three kingdoms of life survey the stability of their tRNA-type substrates and selectively add either CCA or CCACCA. For example, many bona fide tRNAs that are destabilized through mutations or the lack of proper modifications are subjected to CCACCA addition (Wilusz et al., 2011). The CCA-adding enzyme is thus not only critical for generating functional tRNAs, but also likely plays a universal and central role in tRNA and tRNA-like small RNA surveillance and quality control.

Here, we investigated how the CCA-adding enzyme distinguishes structurally stable from unstable RNAs so as to only mark unstable RNAs with CCACCA. We also unravel how a

second CCA cycle can occur despite the enzyme possessing a unique mechanism that normally ensures a single cycle of CCA synthesis. We find that after the first CCA cycle, nucleotide binding to the active site induces the enzyme to apply torque on the RNA. A clockwise screw motion of the enzyme's catalytic domain leads to RNA compression and overwinding. This causes unstable RNAs to extrude a distinctively positioned bulge from the acceptor stem while still bound to the enzyme. The bulge does not perturb the double-helical nature of the substrate and all other structural determinants near the active site remain in place. The catalytic mechanism is thus preserved between the first and second cycles of CCA addition (Pan et al., 2010; Tomita et al., 2006; Xiong and Steitz, 2004). In total, we find that tandem CCA addition is not the result of a modified enzymatic activity that is particular to unstable RNAs. Rather, it is a consequence of the natural activity of the CCA-adding enzyme on a substrate with increased conformational flexibility. By exploiting the versatility of RNA structure, the CCA-adding enzyme is able to trigger the degradation of potentially detrimental small RNAs and tRNAs.

## RESULTS

### Catalysis Is Crucial for the CCA-Adding Enzyme to Detect Unstable RNAs

Certain full-length tRNAs that are hypomodified or contain mutations that destabilize their acceptor stems have previously been shown to be subjected to CCACCA addition (Wilusz et al., 2011). Using the CCA-adding enzyme from *Archaeoglobus fulgidus* (AfCCA), we found that tRNA minihelices, which contain only the acceptor stem and T $\Psi$ C stem-loop, are also efficiently subjected to CCACCA addition when they have guanosines at the first and second positions as well as a destabilized acceptor stem by virtue of mismatches and G-U wobbles (Figure S1A available online). To understand how the CCA-adding enzyme determines whether a substrate is to receive CCA or CCACCA, we first used an arginyl-tRNA<sup>TCG</sup> minihelix that contained a C-A mismatch (C3-A70) in its acceptor stem due to a mutation at position 70 (Figure 1A). This unstable tRNA minihelix (miniUR) was co-crystallized with AfCCA and the structure of the complex was determined by molecular replacement and refined to 2.95 Å resolution (Table S1).

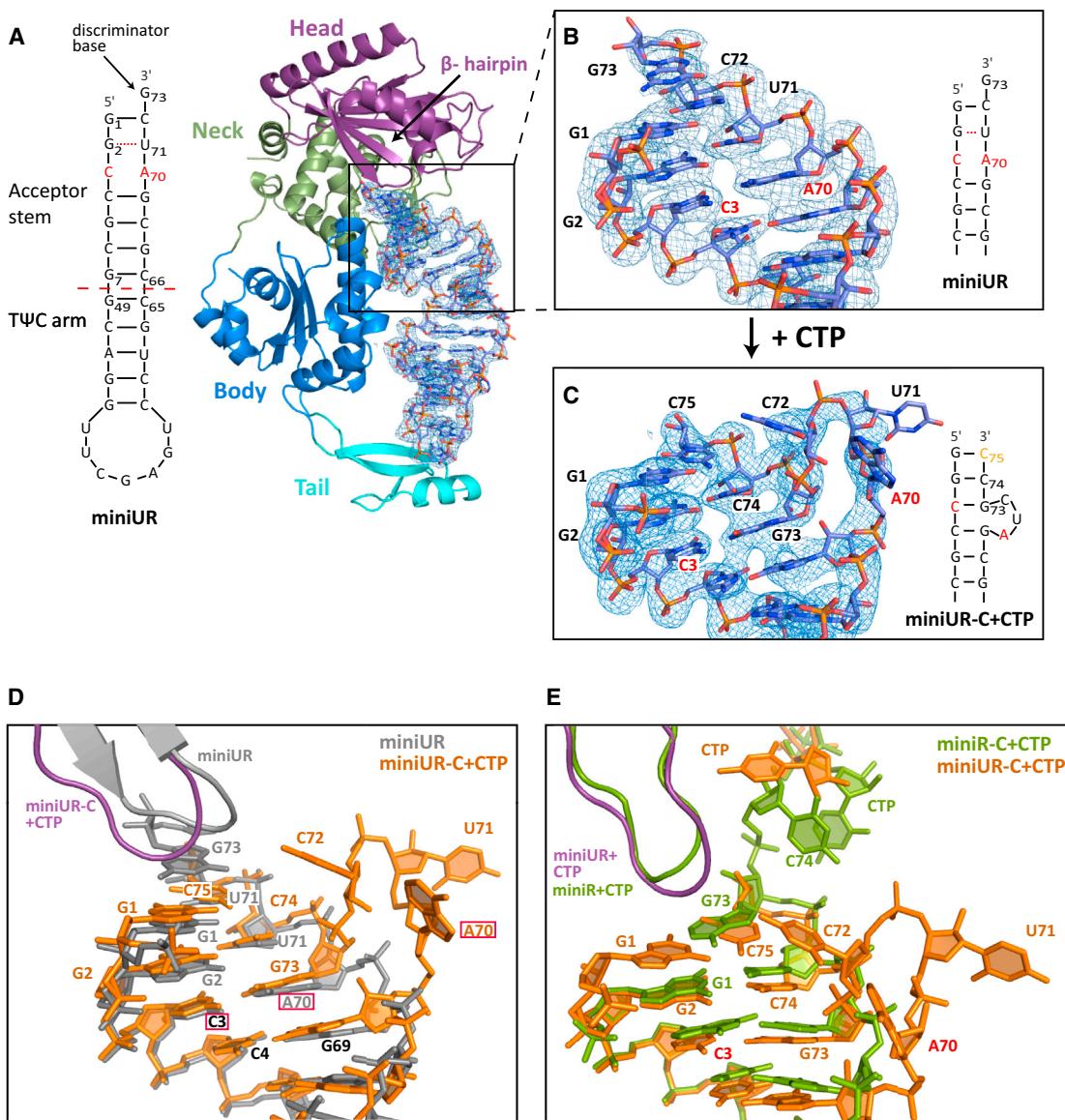
The unstable minihelix is bound between the enzyme's catalytic center, comprised of the head and neck domains, and its tail domain, which serves as a ruler to ensure that only tRNAs and tRNA-like transcripts are substrates for the enzyme (Figure 1A) (Tomita et al., 2004; Xiong and Steitz, 2004). As observed previously, the discriminator base (G73) at the RNA 3'-end extends the stacking of the minihelix and is not rotated out of the helical trajectory (Tomita et al., 2006). However, in contrast to Tomita et al. (2006), which described helical distortion in the middle of the minihelix, we find that the unstable tRNA minihelix perfectly mimics full-length tRNA with its acceptor and T $\Psi$ C stems folding into a continuous A-type RNA helix. Furthermore, the T $\Psi$ C loop is in the same conformation as in full-length tRNA, likely because two nucleotide fragments (derived from the mother liquor after nucleolytic cleavage) aid proper folding by mimicking nucleotides 17 and 18 of full-length arginyl-tRNA<sup>TCG</sup>

(Figure S1B). The C3-A70 mismatch is seamlessly incorporated into the RNA helix and adenine and cytosine face each other, albeit further apart, despite their hydrogen bonding incompatibility (Figure 1B). This larger distance is thus not sufficient to noticeably distort the bound minihelix and, therefore, unlikely to distort a full-length tRNA.

To understand at which step the enzyme is able to detect unstable RNAs, we added CTP prior to co-crystallization. The enzyme catalyzed the addition of two C-nucleotides to the arginyl-tRNA<sup>TCG</sup> minihelix, but diffraction was poor. Nevertheless, we found that using a minihelix already containing the first C nucleotide (C74) improved diffraction to 3.2 Å (miniUR-C+CTP) (Table S1). In this case, the enzyme catalyzed the incorporation of one additional C-nucleotide, resulting in a minihelix ending in -CC. Strikingly, the addition of CTP resulted in a completely remodeled tRNA acceptor stem (Figures 1C and 1D). The original discriminator base G73 moved down by three positions to base pair with C3, resulting in a three-nucleotide bulge (A70-C72) protruding from the RNA helix. Apart from this bulge, which includes the originally mismatched (A70) and wobble (U71) bases, the RNA conformation was virtually unchanged. Two new base pairs G1:C75 and G2:C74 were formed, resulting in a blunt end close to the enzyme (Figure 1C). To determine whether full-length tRNA is able to form a bulge in the same position, we superimposed full-length tRNA (Protein Data Bank [PDB] code 1S21) (Xiong and Steitz, 2004) onto the miniUR-C+CTP structure (Figure S1C). The RNA bulge fits in the space between the enzyme and the anticodon arm, strongly suggesting that an unstable full-length tRNA can form an RNA bulge in the same position. Unfortunately, in the miniUR-C+CTP structure, the enzyme was captured in a catalytically impaired state, as the incoming nucleotide is not properly positioned in the active site. This resembles a complex of the enzyme with either nucleotide alone or with no RNA substrate (Figure S1D) (Tomita et al., 2006; Xiong et al., 2003). Further addition of ATP to our co-crystallization experiments did not yield an actively engaged complex. Although CTP addition results in a catalytically competent preinsertion complex when a stable substrate (denoted miniR) is used (PDB code 2DR5) (Tomita et al., 2006), we observed dramatic refolding of the RNA after CTP addition to an unstable substrate (Figure 1E). This clearly demonstrates that catalysis is necessary for the CCA-adding enzyme to detect unstable RNAs.

### Clockwise RNA Overwinding and Compression upon Active Site Closure

As we were unable to study the second CCA addition cycle by simply supplying nucleotides to the enzyme bound to unstable RNAs (Figures 1C and 1D), we incorporated CCAC at the 3'-end of our unstable model tRNA minihelix (miniUR-CCAC). This RNA was predicted to fold into a single species with a 3nt bulge (Zuker, 2003), reminiscent of our miniUR-C+CTP structure (Figure 1C). In the absence of CTP, triclinic, plate-like crystals formed, and intensity statistics confirmed pseudo-merohedral twinning, accounted for by twin refinement in PHENIX (Adams et al., 2010). Addition of CTP (miniUR-CCAC+CTP) eliminated twinning and the crystals diffracted to 3.15 Å (Table S1). The entire RNA is visible in both structures (Figures 2A and S2A).



## Figure 1. Catalysis Is Crucial to Detect Unstable RNAs

(A) G70 of the arginyl-tRNA<sup>TCG</sup> minihelix was mutated to A, resulting in the miniUR transcript that has significant structural instability within its acceptor stem. The archaeal CCA-adding enzyme is shown in cartoon representation with the head domain in purple, the neck domain in green, the body domain in blue and the tail domain in cyan. Bound RNA is shown in atom colors with oxygens in red, nitrogens in blue, phosphorous in orange and carbons in light blue. 2Fo-Fc density for the bound unstable RNA is contoured at 1  $\sigma$  throughout Figure 1.

**(B) Close-up view of the acceptor stem.**

(C) Close-up view of the acceptor stem following CTP addition (miniUR-C+CTP).

(D) MiniUR before and after CTP addition. Nucleotides labeled in black have similar positions in both structures, the mutation is boxed in red.

(E) Comparison of the CCA-adding enzyme in complex with a wild-type tRNA minihelix (miniR) (PDB code 2DR5) ([Tomita et al., 2006](#)) or in complex with the unstable miniLR minihelix after CTP addition.

See also Figure S1

Prior to nucleotide binding (miniUR-CCAC/open complex), the CCA-adding enzyme adopts an open, inactive conformation consistent with what we observed for the miniUR complex ([Figure 1A](#)) and what is seen for stable RNAs ([Tomita et al., 2006](#)). Confirming the RNA folding prediction, a three nucleotide bulge (nucleotides 70–72) protrudes from the helix and does so in the

same location as seen after CTP addition to the unstable RNA (mini-IR-C+CTP) (**Figures 1C** and **1D**).

CTP binding to the active site (miniUR-CCAC+CTP/preinserter complex) leads to closure of the head domain over the RNA double helix (**Figure 2A**), a process that was previously likened to the closure of a clam shell (**Tomita et al., 2006**). This

leads to compression of the RNA by an entire base-pair step such that G1 of the preinsertion complex is located in the same position as G2 of the open complex (miniUR-CCAC) (Figure 2B; Movie S1). The average helical rise from G1:C72 through C48:G64 decreases from 3.2 Å to 3 Å, accompanied by RNA overwinding of 3° per base-pair step (Movies S1 and S2). Unlike a clam shell, our data suggest that active site closure exhibits a clockwise screw motion of the entire head domain, leading to RNA overwinding (Figure 2C; Movie S2). A similar degree of overwinding was also observed when we reanalyzed active site closure over stable tRNA minihelices (Tomita et al., 2006). Although significant compression occurs in the acceptor stem, the TΨC arm of the minihelix is rather strikingly fixed and binds the tail domain of the enzyme in a similar manner before and after CTP addition (Figure 2A).

Upon comparing the location of the RNA bulges in the cytosine preinsertion complex (miniUR-CCAC+CTP, Figure 2) with two different adenosine preinsertion complexes (miniUR-CCACC+AMPcPP and miniUR-CCACC+CTP, Figure S2B), we found that the RNA was present in by and large the same conformation regardless of how many nucleotides were added at the 3' end (Figure 2D). However, the bulges exhibited varying degrees of disorder. A surface loop in the head domain (Loop1, residues 118–126) stabilizes nucleotide 72 (C72) through a weak hydrogen bond between N4 of C72 and the carbonyl oxygen of Lys124. A second surface loop (Loop2, residues 215–223) in the enzyme's neck domain is oriented toward, but does not bind nucleotide 71. As Loop 1 only interacts with the RNA when it is present in the bulged conformation, we reasoned that inserting additional residues into Loop1 may sterically block bulge formation (thereby blocking CCACCA addition), but not affect the first CCA addition cycle. Three different insertions into Loop1 were tested: (1) insertion of six negatively charged glutamates that are predicted to repel the bulged RNA phosphates close to the enzyme, (2) insertion of three glutamates, and (3) insertion of three large hydrophobic residues (Phe-Leu-Trp) (Figure S2C). Although inserting the hydrophobic residues (Loop1-3Hyd) had no effect on tandem CCA addition, insertion of negatively charged glutamates (Loop1-6Glu and Loop1-3Glu) completely blocked the second CCA-addition cycle, whereas the first CCA triplet was readily added (Figure 2E). As our model arginyl-tRNA<sup>TCG</sup> minihelix contains the wobble base pair G2:U72, which may disproportionately favor refolding, we confirmed that two additional minihelices lacking this wobble pair gave identical results (Figure S2D).

We conclude that bulge formation is critical for tandem CCA addition and RNA surveillance of unstable RNAs. Furthermore, RNA compression and overwinding, induced by the head domain acting as a "molecular screwdriver," appears to be a way for the CCA-adding enzyme to select for proper substrates by challenging the stability of the substrates.

### Menβ tRNA-like Small RNA Behaves Like an Unstable tRNA and Is Held onto the CCA-Adding Enzyme through Ionic "Tweezers"

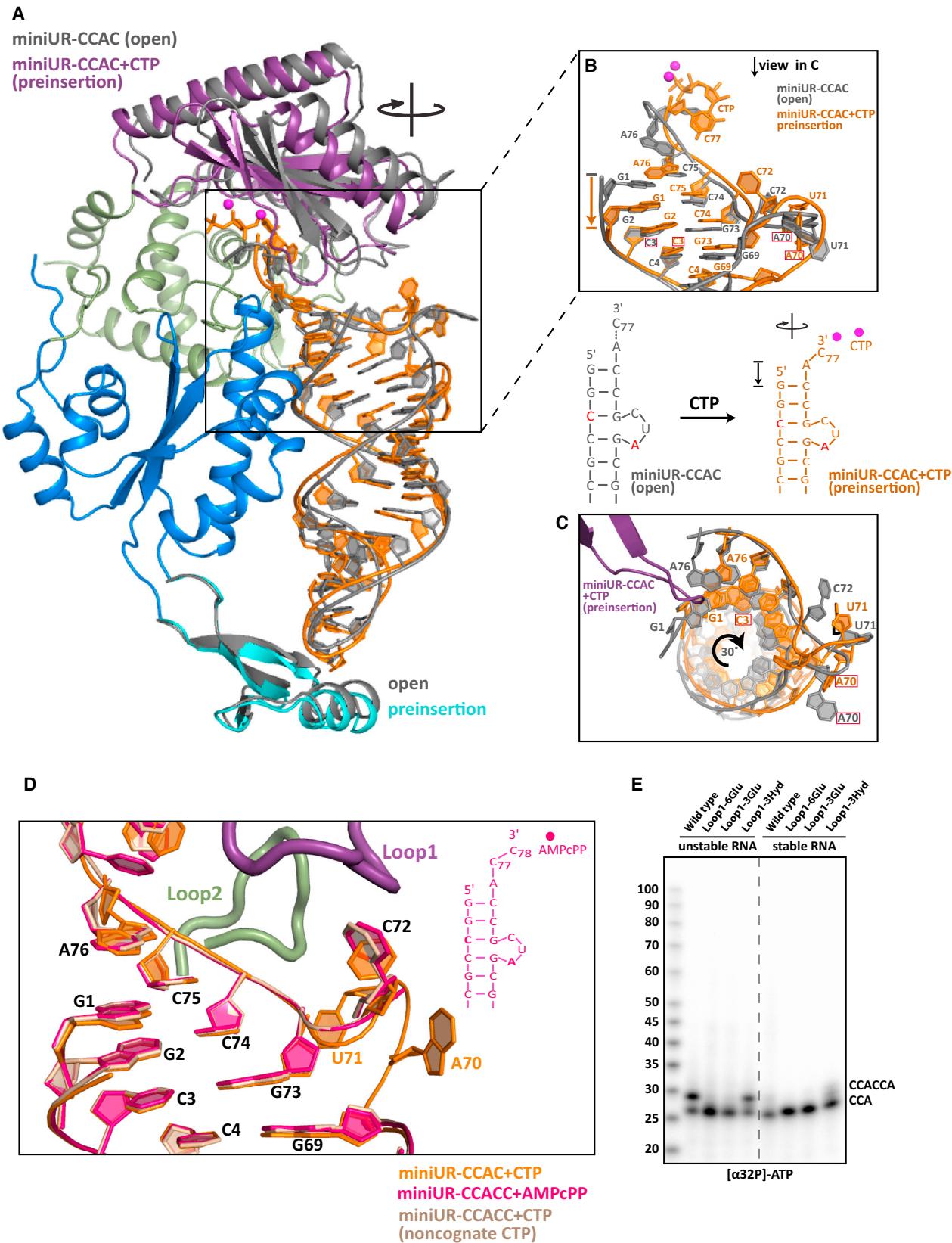
Upon comparing cytosine selection and incorporation during the first and second CCA cycle, we noticed a strict conservation of the catalytic mechanism between the two cycles (Figure S3A). Since the same holds true for adenosine incorporation (Fig-

ure S2B), we were surprised to observe a kinked minihelix in the adenosine preinsertion complex of AfCCA bound to a human Menβ-derived minihelix (miniMβ-CCACC+AMPcPP) (Figures 3A and S3B). Despite diffraction to 2.6 Å (Table S1), the entire 3nt RNA bulge and the adjacent base U54 are invisible. In addition, G4 shows two distinct conformations (Figures 3A and 3B) at the center of a clear 11° kink in the helical axis of the Menβ minihelix (Figure 3B). Nevertheless, the two ends of the Menβ minihelix, the three 5'-terminal "acceptor-stem" base pairs and the "TΨC-arm" base pairs, superimpose perfectly with the equivalent features of the unstable arginyl-tRNA<sup>TCG</sup> minihelix structure (miniUR-CCACC+AMPcPP) (Figure 3B). Since the kink essentially uncouples the acceptor stem from the TΨC arm, it seemed that the tail domain may not actually be anchoring the RNA on the enzyme, as has been previously proposed (Cho and Weiner, 2004; Tomita et al., 2006), during the second cycle of CCA addition.

To gain insights into whether a kinked Menβ tRNA-like substrate conforms to the normal CCA-adding enzyme mechanism, we plotted the temperature-factor distribution of the entire mini-Mβ-CCACC+AMPcPP complex (Figures 3C and S3C). Strikingly, the most stable region of the entire complex is the active site cleft and the adjacent top part of the acceptor stem. The tail domain and the TΨC arm, on the other hand, are highly mobile. Analogous results were obtained for canonical tRNA minihelices undergoing the second CCA cycle (miniUR-CCACC+AMPcPP and miniUR-CCACC+CTP), which all exhibit identical crystal packing (Figure S3D). In contrast, the tail domain and the TΨC arm are both stable during the first CCA addition cycle (Figure S3E). Therefore, it appears that the tail domain does not serve as the major RNA anchor during the second CCA addition cycle.

Why then is the top of the acceptor stem so stable? Upon calculating the electrostatic surface potential of the CCA-adding enzyme, we noticed two positively charged surface patches binding the top of the acceptor stem (Figure 3D). Patch1, consisting of body domain residues Gln296, Arg299, Arg302, and Lys402, binds the 5'-end of Menβ via direct ionic interactions to bridging RNA phosphates. Patch2 is comprised of Arg129 from the head domain and the catalytically essential Arg224 from the neck domain. Interestingly, most contacts on this 3'-side are water-mediated and, therefore, weaker than on the 5'-side (Figure 3D). To address whether those ionic interactions keep the unstable RNA in place for the second CCA-addition cycle, we mutated both Arg299 and Arg302 to alanine and tested enzymatic activity in vitro (Figure 3E). While the wild-type enzyme was able to extend the unstable arginyl-tRNA<sup>TCG</sup> minihelix to CCACCA at a nucleotide concentration of 0.5–5 μM, the mutant enzyme required 10–100× higher concentration before switching from CCA to CCACCA addition (Figure 3E). These results are consistent with Patches1+2 forming "ionic tweezers" (Figure 3D, schematic) that ensure the second CCA cycle proceeds efficiently.

Apart from a kink in the acceptor stem helix, we conclude that tandem CCA addition proceeds in a highly similar fashion for canonical tRNAs and tRNA-like small RNAs. Furthermore, it is clear that the tail domain of the CCA-adding enzyme has two separate functions. First, it serves as a ruler to ensure selective binding of



(legend on next page)

tRNA-type substrates. Second, it allows torque to be applied on the helix upon active site closure and thereby contributes to the discrimination between stable and unstable RNA substrates.

### On-Enzyme RNA Refolding Permits the Second CCA Addition Cycle

As the CCA-adding enzyme is renowned for its precision in terminating synthesis after a single CCA triplet, we were eager to determine how it transitions from the first to the second cycle of CCA addition. For unstable RNAs, we hypothesized that this transition may occur in one of two ways: (1) the RNA may dissociate from the enzyme after the first round of catalysis and re-bind in the bulged conformation to initiate the second cycle or, alternatively, (2) the RNA may refold while bound to the enzyme.

Initial support for the latter model came from determining the structure of AfCCA in complex with the mutant arginyl-tRNA<sup>TCG</sup> minihelix ending in CCACCA that was bromouridine (<sup>Br</sup>U)-labeled at nucleotides 54, 63, and 71 (miniUR-CCACCA) (Figure 4A). Two strong anomalous peaks were detected at the expected positions for <sup>Br</sup>U54 and <sup>Br</sup>U63 (Figure 4A). Surprisingly, we also observed a weak anomalous peak for <sup>Br</sup>U71 near the top of the acceptor stem. This result was unexpected since U71 falls within the location of the bulge, which was disordered in this structure. The only reasonable explanation is the presence of two alternative RNA conformations, one representing the RNA prior, and one after refolding. Indeed, we were able to refine the non-refolded state of the RNA to an occupancy of 30%, thereby providing an explanation for the weak anomalous peak. The remaining 70% is disordered in the bulge. Considering that the electron density for the tRNA 5'-end was unambiguous in this structure (as in all other structures of this study) and that the tRNA 5'-end is tightly bound by the enzyme (Figures 3C–3E), we hypothesized that the tRNA 3'-end refolded independently of the 5'-end to allow addition of the second CCA.

To test this model, we used two different approaches: first by covalently tethering the RNA to the enzyme, and second using a competition experiment. For the first, we were inspired by earlier work on DNA (Huang et al., 1998) and we set out to reversibly crosslink the CCA-adding enzyme to the unstable tRNA using a site-specific link. We reasoned that if tandem CCA addition proceeded efficiently when the RNA substrate and CCA-adding enzyme were covalently joined, refolding most likely occurred “on-enzyme.”

To engineer a covalent complex with a site-specific linkage, a purine nucleoside analog bearing a short linker terminating in a free thiol was incorporated at position 56 of the TΨC arm of stable (WT) and unstable (mutant) arginyl-tRNA<sup>TCG</sup> minihelices (Figure S4A) (Peacock et al., 2011). The TΨC arm seemed to be the best location for not interfering with catalysis, as much of the rest of the minihelix is compressed prior to catalysis (this study and Tomita et al., 2006). Likewise, Asp351 and Arg344 of the AfCCA tail domain are located closest to the crosslinkable nucleotide and were therefore replaced with cysteines to allow a protein-RNA disulfide crosslink to form, a prediction we confirmed experimentally (Figure S4B).

Since tRNA surveillance through CCACCA addition is conserved across all kingdoms of life (Wilusz et al., 2011), we also wanted to include class II CCA-adding enzymes. Given a lack of sufficient structural information on class II enzymes, the structure of the full-length human mitochondrial CCA-adding enzyme was determined to 1.9 Å (Figure S4C). Although the tail domain was disordered in Augustin et al. (2003), it is clearly resolved in our structure, and we were able to construct a model of a tRNA-bound complex (Figure 4C). Based on this model and allowing for conformational changes in the protein, several cysteine mutants were tested for their ability to form disulfide bonds with the crosslinkable tRNA. Two mutants, G364C and G379C, crosslinked efficiently to both stable and unstable tRNA minihelices (Figure 4D).

We next purified these crosslinked complexes to test whether RNA refolding occurs on the enzyme in a standard CCA addition assay (Shi et al., 1998b). Reactions were performed in the presence of 2 μM cold ATP and CTP, combined with either radioactive ATP or CTP to measure A or C incorporation, respectively. Since the minihelices used ended in -CC (miniUR-CC), further C incorporation served as a sensitive readout of refolding, as it should only occur during the second CCA cycle, after the substrate has successfully completed the first cycle. A stable tRNA minihelix (miniUR-CC) was used as a control as it should terminate synthesis after the first CCA cycle and thus not incorporate CTP.

When we measured radioactive ATP incorporation onto the 3'-end of stable and unstable human complexes, we detected a lower molecular weight product for the stable tRNA complex and a higher molecular weight product for the unstable tRNA complex, as would be expected for CCA versus CCACCA addition (Figure 4E, lanes 1 and 2). Surprisingly, similar results were

**Figure 2. Clockwise Screw Motion and RNA Compression upon Active Site Closure**

Structural transitions upon active site closure.

(A) RNA and protein domains of the open complex (miniUR-CCAC) are shown in cartoon representation in gray. The preinsertion complex (miniUR-CCAC+CTP) is colored as in Figure 1A with miniUR-CCAC+CTP RNA, including the incoming CTP, in orange. Catalytic magnesium ions are in magenta. The clockwise screw motion of the head domain and RNA is indicated.

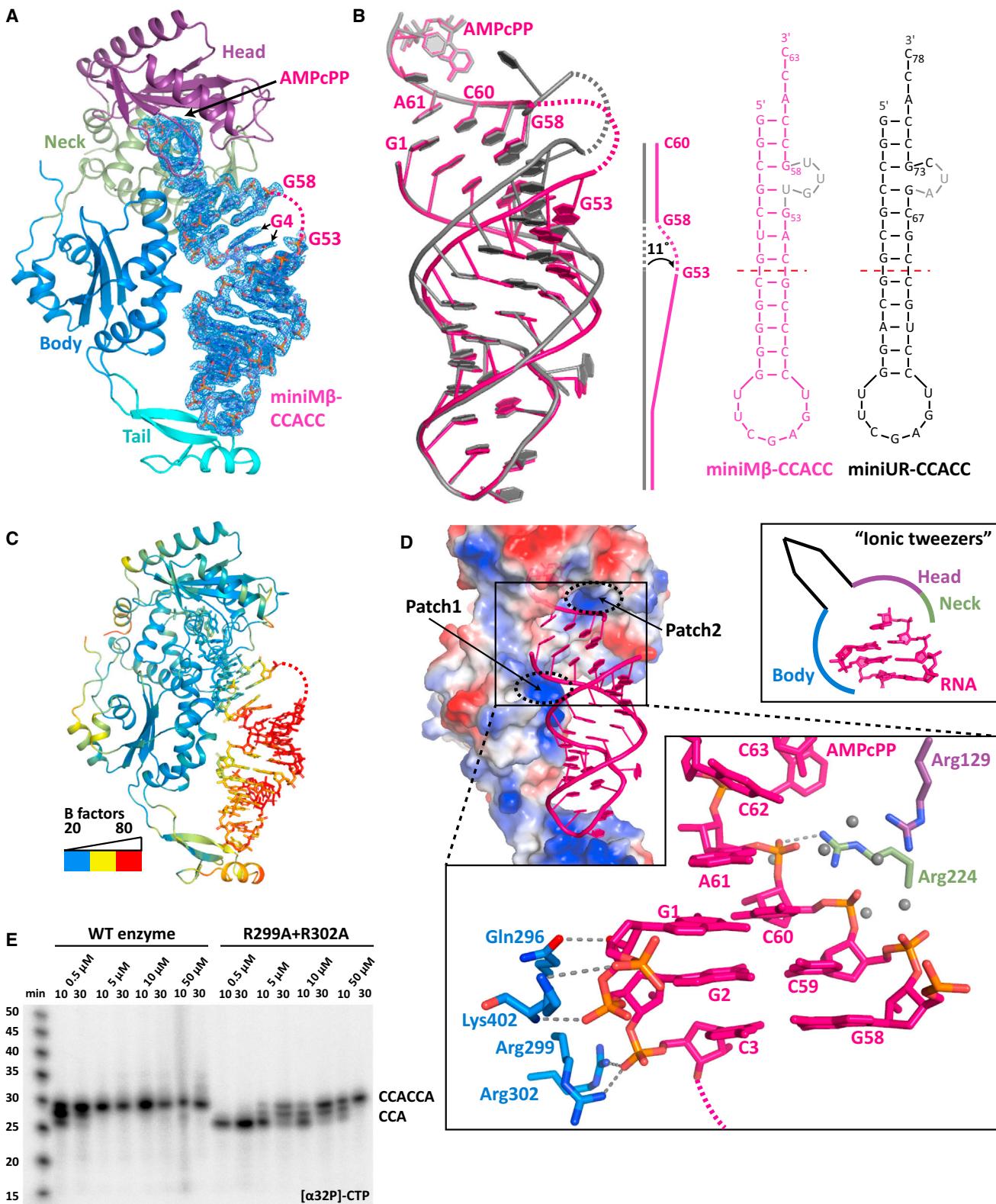
(B) Close-up view of the acceptor stem of miniUR-CCAC (gray) and miniUR-CCAC+CTP (orange). The A70 mutation is boxed in red. RNA compression is indicated by an arrow.

(C) Top-down view of (B) including the miniUR-CCAC+CTP β-hairpin (purple). Clockwise RNA rotation is indicated.

(D) Location of the RNA bulge in all three preinsertion complex structures of the second CCA cycle. Common nucleotides to all structures are numbered in black, unique ones follow the RNA color scheme. Loops 1 and 2 are shown for miniUR-CCAC+CTP only.

(E) In vitro CCA-addition assays were performed to assay the ability of AfCCA containing Loop 1 insertions to incorporate [ $\alpha$ -<sup>32</sup>P] ATP onto stable or unstable (G70A) arginyl-tRNA<sup>TCG</sup> minihelices ending in -C. The insertion of three or six glutamates abolished the ability of the CCA-adding enzyme to mark unstable RNAs for degradation by adding CCACCA.

See also Figure S2 and Movies S1 and S2.



**Figure 3. Unstable RNA Is Held onto the CCA-Adding Enzyme through Ionic “Tweezers”**

(A) Structure of the adenosine preinsertion complex of the Men $\beta$  tRNA-like small RNA (miniM $\beta$ -CCACC+AMPcPP), as in Figure 1A. The disordered RNA bulge is dashed and colored in pink.

(legend continued on next page)

obtained using radioactive CTP (Figure 4E, lanes 3 and 4), despite our expectation not to detect CTP incorporation into stable tRNAs ending in -CC. Upon reversing the crosslink by adding 20 mM DTT prior to performing the CCA addition assay, CTP incorporation was detected only onto the unstable tRNA, indicating that the stable tRNA now completed only the first CCA cycle (Figure 4E, lanes 7 and 8). Interestingly, all RNA products were smaller under reducing (Figure 4E, lanes 5–8) versus oxidizing conditions (Figure 4E, lanes 1–4), with the products under reducing conditions running at the same sizes as never-cross-linked reaction controls (Figure 4E, lanes 9 and 10). We therefore conclude that crosslinking the human CCA-adding enzyme to its RNA substrate resulted in tandem CCA addition to stable RNAs and three CCA triplets being added to unstable RNAs. Similar results were obtained with the crosslinked archaeal enzyme, although the difference between stable and unstable RNA was less pronounced (Figure S4D). We should note that upon crosslinking the substrate to either enzyme and increasing the nucleotide concentration to 50 μM, both enzyme classes behaved like “oligo-CCA”-adding enzymes and added extended CCA tails to both stable and unstable RNAs, with unstable RNAs harboring longer tails on average (Figure S4E). This “oligo-CCA” addition is not observed in vivo (Wilusz et al., 2011) since refolding and CCA addition rates are much higher for crosslinked complexes compared to those of the uncrosslinked versions.

For the competition experiment, we reasoned that if the RNA refolded between the two CCA-addition cycles while bound to the enzyme, a pre-bound tRNA would not be competed off the enzyme by another RNA when transitioning from the first to the second CCA cycle. We, therefore, pre-bound unstable full-length arginyl-tRNA<sup>TCG</sup> (G70A) ending in its discriminator base to either the human or the archaeal CCA-adding enzyme in the presence of CTP only. After having incorporated two C nucleotides, we expected the enzymes to remain bound to their substrate RNAs, awaiting ATP to complete the first CCA cycle. ATP was then added simultaneously with an excess of full-length human Menβ tRNA-like small RNA already ending in CCA, and product formation was monitored. Consistent with the tRNA refolding on enzyme, pre-binding of unstable arginyl-tRNA<sup>TCG</sup> to both classes of enzymes rendered the enzymes “immune” toward excess RNA and resulted in similar amounts of mature arginyl-tRNA<sup>TCG</sup> product, irrespective of the competing Menβ concentration (Figure 4F, lanes 1–4). In contrast, when both arginyl-tRNA<sup>TCG</sup> and Menβ substrates ending in -CC were added at the same time, final products were predominantly synthesized for Menβ, the substrate present in excess (Figure 4F, lanes 5–8). We note that human Menβ is a better substrate for the human CCA-adding enzyme compared to the archaeal enzyme (Fig-

ure 4F). Also, when the archaeal reaction scheme was inverted such that human Menβ was pre-bound and arginyl-tRNA<sup>TCG</sup> served as the competitor, we observed a less pronounced trend, since Menβ is not as good a substrate as arginyl-tRNA<sup>TCG</sup> (Figure S4F).

Taken together, it is clear that tandem CCA addition does not involve RNA dissociation. Instead, RNA refolding takes place “on-enzyme.”

### Active Site Closure Gauges RNA Stability and Triggers Refolding

What triggers “on-enzyme” RNA refolding to enable the second CCA cycle? Termination of the first CCA cycle, which precedes refolding, has previously proven difficult to study. It is, however, known that in the termination state, the entire CCA triplet stacks onto the acceptor stem and abuts against the β sheets containing the active site residues, thereby blocking the addition of further nucleotides (Xiong and Steitz, 2004).

Insight regarding how refolding is triggered was gained from a 2.7 Å structure of a monoclinic crystal of the Menβ tRNA-like small RNA adenosine preinsertion complex (miniMβ-CCACC+AMPcPP-i). Most of the RNA bound in the active site pocket of this complex is ordered, despite the pocket not being entirely closed (Figures 5A and S5A). As expected for a substrate ending in CCACC, the tRNA-like transcript is bound in its refolded state, ready for the second CCA cycle to be completed with the addition of an A. Nucleotides C59 and C60 (equivalent to C74 and C75 for tRNAs) form strong base pairs to G1 and G2. A61 (A76 for tRNAs), the last A of the first CCA triplet, is in a similar position to a discriminator base. However, its base edge projects toward a β-hairpin that has been previously shown to be critical for catalysis and proofreading (Cho et al., 2005). This A also stacks with His97, thereby displacing Asp96 and Ala95. C62 (C77 for tRNAs), the first C of the second cycle, packs against Tyr99, whereas C63 (C78) folds back onto the 3'-end of the acceptor stem, lying orthogonal to it and tucking against Arg129 (Figures 5A and S5A). Although electron density for the C63 base is weak, its 3'-hydroxyl group is clearly not in a position to attack the α-phosphate of the incoming nucleotide. Interestingly, the incoming AMPcPP is also not properly locked into its binding pocket as its base edge is not recognized by Arg224, nor do the three phosphate groups coordinate metal ion A.

When we compared this structure to the proper adenosine preinsertion complexes of Menβ (miniMβ-CCACC+AMPcPP, Figure 3A) and unstable tRNA (miniUR-CCACC, Figure S2B), we realized that the observed state represents a functional intermediate between the open (inactive) and the closed preinsertion complexes. The bound small RNA exhibits an intermediate

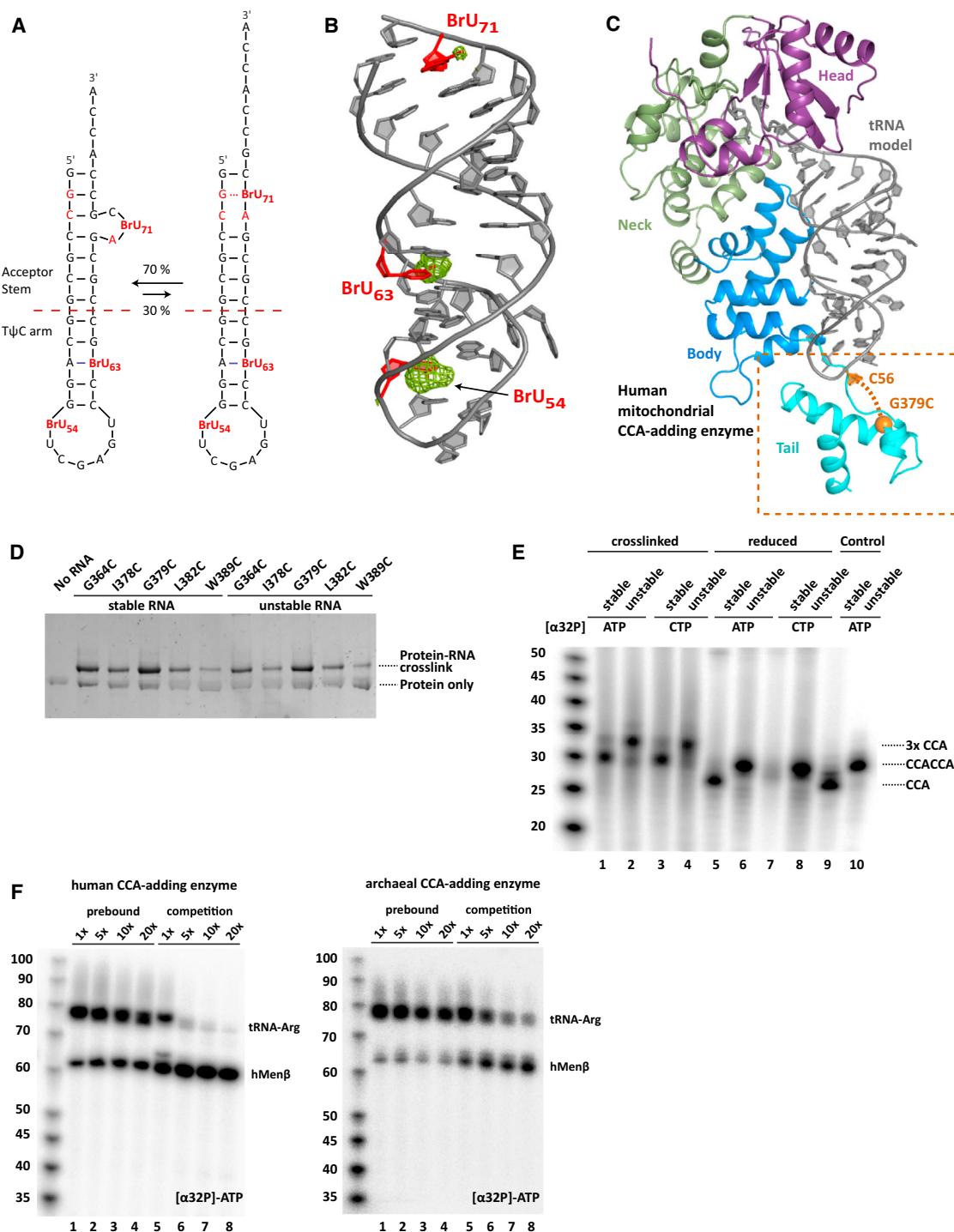
(B) Comparison between the miniMβ-CCACC+AMPcPP and miniUR-CCACC+AMPcPP RNA structures. Both RNAs are schematically diagrammed on the right with disordered nucleotides in gray. The kink in miniMβ-CCACC RNA is sketched.

(C) Temperature factor distribution of miniMβ-CCACC+AMPcPP. Low temperature factors of 20 Å<sup>2</sup> are in blue, intermediate values in yellow, and values above 80 Å<sup>2</sup> in red.

(D) Electrostatic surface potential of the archaeal CCA-adding enzyme. Blue depicts positively charged, white neutral, and red negatively charged areas. The two positively charged patches holding the acceptor stem are indicated. MiniMβ RNA is in cartoon and colored in pink. The close-up view details the ionic interactions that glue the top of the acceptor stem to the enzyme. A schematic of the “ionic tweezers” is shown.

(E) In vitro CCA-addition assays using wild-type or mutant AfCCA and the unstable (G70A) arginyl-tRNA<sup>TCG</sup> minihelix (miniUR) substrate. Weakening the strong interactions between the RNA 5'-end and the enzyme impaired the second CCA-addition cycle, whereas the first cycle was unaltered.

See also Figure S3.



**Figure 4. On-Enzyme RNA Refolding Permits the Transition from the First to the Second CCA Cycle**

(A) The two interchangeably bound RNA conformations observed in bromouridine (<sup>Br</sup>U)-labeled miniUR-CCACCA co-crystals. <sup>Br</sup>U and RNA instabilities are highlighted in red.

(B) The double-helical part of the non-refolded RNA from (A) is shown in cartoon representation and colored in gray. <sup>Br</sup>U are in red. The anomalous difference Fourier map is colored in green and contoured at 3.5 σ.

(C) Structure of the full-length human mitochondrial CCA-adding enzyme. Protein domains are colored according to the archaeal enzyme in Figure 1A. The newly resolved tail domain is highlighted. A tRNA minihelix is modeled according to PDB code 1VFG (Tomita et al., 2004). An ethylthio-purine crosslinker was incorporated at tRNA position C56 and crosslinked to the CCA-adding enzyme carrying a G379C mutation.

(legend continued on next page)

degree of compression and the  $\beta$ -hairpin displays an intermediate degree of closure (Figure S5B). Similar to our observations for unstable tRNA (Figure 2), the transition from the intermediate to the preinsertion complex revealed a coordinated clockwise screw motion of the head domain, resulting in RNA overwinding and compression of the acceptor stem by an entire base pair step (Figure 5B; Movies S3 and S4). The head domain pushed the protruding nucleotide A61 of the intermediate state back into its proper position to stack onto the 3'-end of the acceptor helix, unstacking it from His97. The  $\beta$ -hairpin is now in a proper conformation to proofread C62, which undergoes a 90° rotation, whereas C63 reorients its 3'-hydroxyl toward Asp110, which serves as a general base during adenosine incorporation (Pan et al., 2010). The incoming AMPcPP moves to its proper preinsertion position, interacting with Arg224 and coordinating Metal B along with the catalytic residues Glu59 and Asp61. Interestingly, only the Men $\beta$  preinsertion complex shows a kinked helix, likely induced by complete active site closure (Figures 3A and S5C).

The apparent correlation between active site closure and RNA compression indicates that apart from being essential for catalysis, active site closure is the way by which the stability of the RNA is assessed. Following the first CCA addition, the torque applied on the RNA double helix essentially gauges weak spots in the duplex and causes refolding of the RNA in the form of extrusion of a bulge.

### Nucleotide Binding Is Critical for Triggering RNA Substrate Refolding

Three lines of evidence suggested that nucleotide binding is critical to trigger refolding. First, our ability to isolate an intermediate complex is likely attributed to the presence of a sub-optimally positioned nucleotide in the active site pocket. Proper binding of the triphosphate moiety locks the enzyme in the catalytically active preinsertion state. Second, the termination state after the first CCA cycle (PDB 1SZ1) (Xiong and Steitz, 2004) has sufficient space for an incoming nucleotide, which could start off the second CCA-addition cycle (Figure 5C). Finally, since active site closure is accompanied by a clockwise rotation, we reasoned that the termination complex is formed from the adenosine preinsertion complex by a counter-clockwise motion of the head domain, which appears to be the case (Figure 5C). This movement is essential to reset the enzyme prior to a new cycle of CCA addition.

As nucleotide binding is critical to trigger refolding, we hypothesized that increasing the nucleotide concentrations in standard CCA-addition assays (Shi et al., 1998b) may lower the threshold of instability needed in a tRNA acceptor stem for CCACCA addition to occur. Using a nucleotide concentration of 0.5  $\mu$ M, we

recapitulated what is seen in vivo for full-length wild-type or mutant arginyl tRNA (Figure 5D). However, at nucleotide concentrations above 5  $\mu$ M, the wild-type tRNA was converted into a tandem CCA target by both classes of CCA-adding enzymes (Figure 5D). Similar results were obtained for minihelix substrates (Figure S5D). At nucleotide concentrations above 20  $\mu$ M, we detected even further extension of the CCA tails, reminiscent of the results obtained with crosslinked complexes (Figure 4E). Most surprisingly, both enzyme classes converted stable tRNAs into CCACCA targets at identical nucleotide concentrations, which was unexpected since tRNA approaches their respective active sites from entirely different directions (Xiong et al., 2003). Whereas the extruded RNA bulge tucks against the protein for class I enzymes (Figure 2), it almost certainly protrudes into the solvent in class II enzymes (Figure S5E).

Additional tRNAs were subsequently tested to determine whether sequence elements within the tRNA may affect the ability of nucleotide concentration to regulate CCA versus CCACCA choice. As in our model arginyl-tRNA<sup>TCG</sup> substrate, wild-type cysteinyl-tRNA<sup>GCA</sup> contains a wobble base pair in its acceptor stem and was converted into a tandem CCA target at a nucleotide concentration of 5  $\mu$ M (Figure S5F). On the other hand, wild-type tRNAs with acceptor stems containing only Watson-Crick base pairs required nucleotide concentrations to be at least ten times higher (50  $\mu$ M) before they were converted into CCACCA targets (Figures S5G and S5H).

We suggest that the choice between CCA versus CCACCA addition is made when a nucleotide binds to the active site in the termination state following the first CCA addition cycle. For stable substrates, clockwise active site closure triggered by nucleotide binding leads to RNA dissociation as the energetic cost of breaking multiple Watson-Crick base pairs is too high. Unstable RNAs, however, will refold "on-enzyme" and a second CCA triplet will be added. Elevated nucleotide concentrations shift the reaction equilibrium toward tandem CCA addition with the actual concentration wherein this switch occurs dependent on acceptor stem stability. Elevated nucleotide concentrations might favor a closed active site, which would trap even transiently refolded stable substrates after the first CCA cycle on the enzyme by adding the first C of the second cycle.

### RNAs with Longer Bulges Are Not Readily Accommodated by the CCA-Adding Enzyme

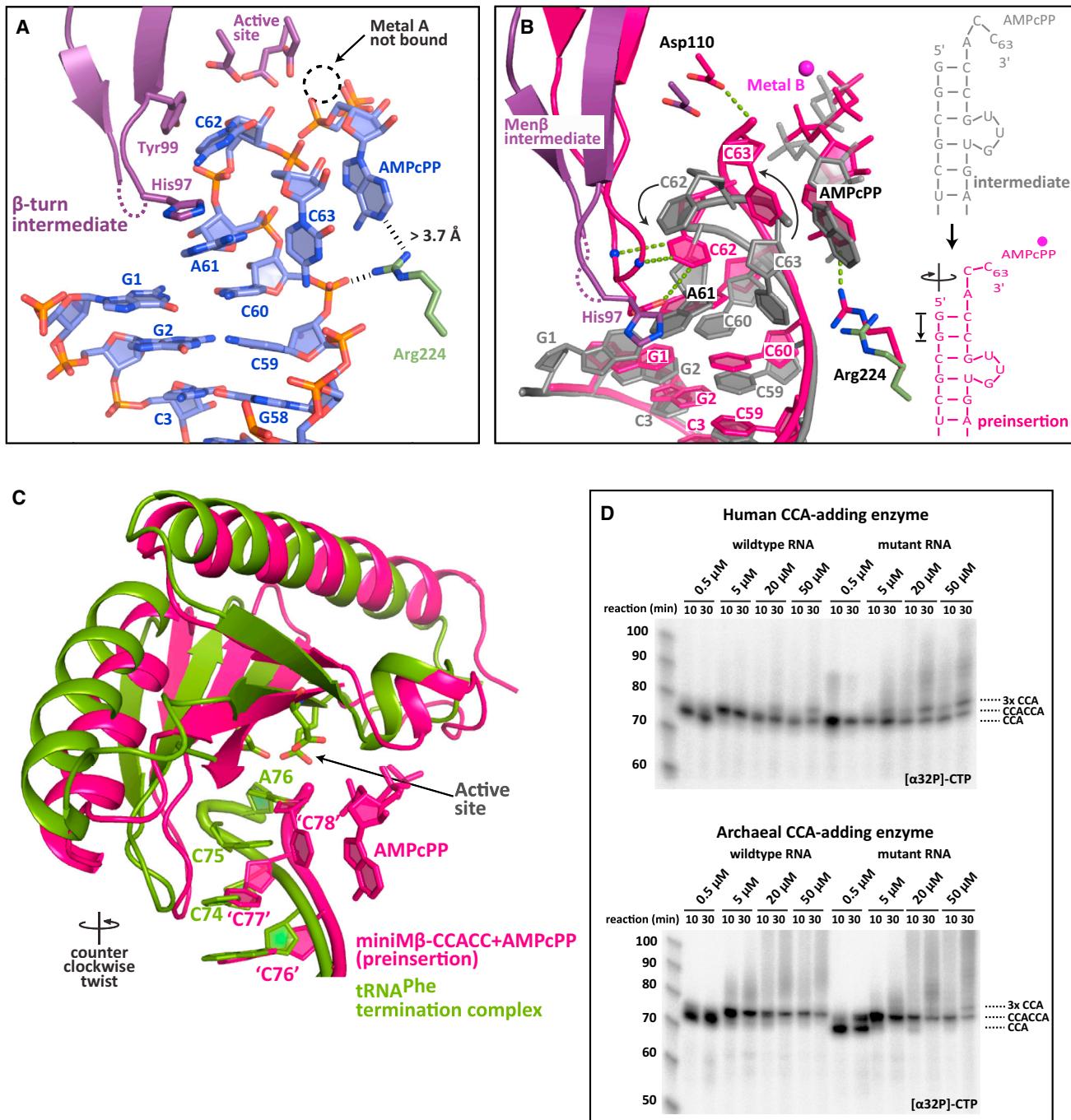
Why, however, is CCA addition limited to a tandem triplet in vivo? Although extended tails longer than CCACCA are efficiently added under certain conditions *in vitro* (higher nucleotide concentrations or when the RNA is covalently bound to the enzyme), they

(D) Silver-stained SDS-PAGE of the crosslinked human CCA-adding enzyme and the ethylthio-purine-derivatized minihelices. Crosslinked complexes show retarded gel mobility, with the G379C and G364C mutants showing the highest crosslinking efficiency.

(E) *In vitro* CCA addition assays with crosslinked complexes. The wild-type and G70A mutant arginyl-tRNA<sup>TCG</sup> minihelices are denoted as stable and unstable, respectively. Reactions 2–5 were under oxidizing, whereas reactions 6–10 were under reducing conditions. Radioactive nucleotides are indicated above the lanes. Unstable minihelices that had never been crosslinked were used as controls (reactions 9–10).

(F) *In vitro* competition experiments. In lanes 1–4, full-length arginyl-tRNA<sup>TCG</sup> (G70A) was pre-bound to the human and archaeal CCA-adding enzyme in the presence of 2  $\mu$ M CTP before adding cold ATP supplemented with [ $\alpha^{32}$ P]-ATP and varying amounts of the hMEN $\beta$  transcript (up to a 20-fold excess). In lanes 5–8, both RNAs were added simultaneously along with ATP. An equal amount of arginyl-tRNA<sup>TCG</sup> was added in all lanes, whereas hMEN $\beta$  amounts varied as denoted at the top. Pre-bound RNA was efficiently extended to CCACCA irrespective of added excess RNA, indicating that the bound RNA refolds "on-enzyme." When added simultaneously, both RNAs compete for binding to the CCA-adding enzyme.

See also Figure S4.



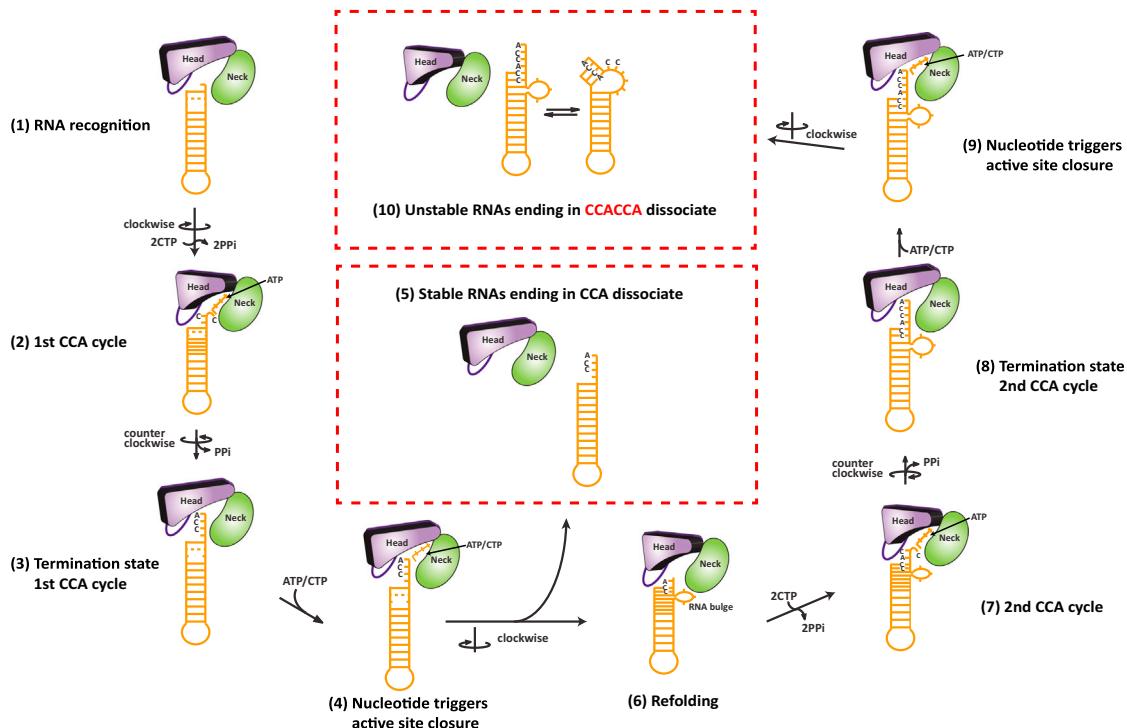
**Figure 5. Active Site Closure Gauges RNA Stability and Triggers Refolding**

(A) The intermediate adenosine preinsertion complex of Men $\beta$  (miniM $\beta$ -CCACC+AMPcPP-i) shown as in Figure 1A. Residues contributed by the head domain are in atom colors with carbons in purple. The disordered residues Ala95 and Asp96 are dashed in purple. Arg224 from the neck domain is in atom colors with carbons in green. A dashed circle highlights the missing Metal B. Improper positioning of the incoming nucleotide is shown with black dashes.

(B) Superposition of the adenosine intermediate and preinsertion complexes (miniM $\beta$ -CCACC+AMPcPP). The intermediate state RNA and AMPcPP are in gray, while the intermediate β-hairpin and Asp110 are in purple. Intermediate Arg224 is in green. The entire preinsertion complex (miniM $\beta$ -CCACC+AMPcPP) is in pink. Three proofreading interactions, the AMPcPP base edge recognition and the Asp110 general base contact to the 3'-hydroxyl are shown with green dashed lines for the preinsertion complex, as is Metal B bound to AMPcPP.

(C) Comparison of the miniM $\beta$ -CCACC+AMPcPP complex and the termination complex of the first CCA cycle (PDB code 1S71) (Xiong and Steitz, 2004). The incoming AMPcPP is depicted for Men $\beta$ . Canonical tRNA numbering is used for Men $\beta$  to simplify comparison. The active site residues that the RNA stacks against

(legend continued on next page)



**Figure 6. Model of RNA Surveillance by the CCA-Adding Enzyme**

The head and neck domains of the CCA-adding enzyme are shown as abstract bodies colored in purple and green, respectively, with the  $\beta$ -hairpin as an oval. Bound unstable RNA is shown in orange with its compression state changing during the CCA-addition cycles. Steps 1–10 are described in the text. See also Figure S6.

are only rarely added in vivo (Wilusz et al., 2011, 2012). To understand why, we carried out fragment-based assembly of RNA with full atom refinement (FARFAR), a ROSETTA-based de novo RNA prediction algorithm (Das and Baker, 2007; Das et al., 2010). Using the miniUR-CCAC+CTP complex as a template (Figure 2), 5,000 structures were calculated assuming either a 3nt (for tandem CCA) or a 6nt bulge (for three CCAs). As expected, models containing a 6nt bulge deviated more significantly from the template structure and none formed a continuous helix (Figure S6A). In contrast, about a third of the models with a 3nt bulge exhibited a continuous helix (Figure S6B).

We posit that in order to apply torque onto the acceptor stem to trigger refolding, a continuous helix would be necessary. From these calculations, it appears that the reason that three CCAs are not accommodated is because the resulting RNA molecules are unable to form continuous helices.

## DISCUSSION

### Model of Tandem CCA Addition

tRNA surveillance through tandem CCA addition is a universally conserved mechanism by which the cell distinguishes between

stable and certain unstable tRNAs and tRNA-like transcripts (Wilusz et al., 2011). While stable tRNAs receive only CCA, unstable tRNAs beginning with GG are marked with CCACCA with near 100% efficiency and rapidly degraded (Wilusz et al., 2011). We show that the structural flexibility of RNA controls the choice between CCA versus CCACCA addition and propose the following model for tandem CCA addition (Figure 6): (1–3) a proper tRNA-type substrate is first recognized by the CCA-adding enzyme. The discriminator base is inserted between the head and neck domains, while the  $T_{\psi}C$  loop interacts with the enzyme's tail domain (Figure 1A). The first CCA addition cycle then proceeds as has been established (Pan et al., 2010; Tomita et al., 2006; Xiong and Steitz, 2004). Following the addition of the terminal A, pyrophosphate is released and the head domain rotates counter-clockwise to allow continuous stacking of the CCA triplet against the head domain (Figure 5C). (4) Nucleotide binding at this stage induces the clockwise closure of the active site with torque applied on the RNA duplex (Figures 2 and 5). This “interrogation” of the RNA by the enzyme results in two possible outcomes: (5) if the bound RNA is stable, it will dissociate from the enzyme since breaking multiple stable Watson-Crick base pairs for RNA refolding is unfavorable; (6) however, if the RNA is

in the termination complex are shown as sticks. The counter-clockwise twist between the Men $\beta$  preinsertion complex and the tRNA termination complex is indicated.

(D) In vitro CCA addition assays with full-length stable and unstable (G70A) arginyl-tRNA<sup>TCG</sup> and either the human or the archaeal CCA-adding enzyme. Cold nucleotide concentrations (ATP+CTP) were increased in the presence of radioactive CTP. CCA tail length is indicated. See also Figure S5 and Movies S3 and S4.

unstable, the screw motion will trigger on-enzyme refolding of the RNA 3'-end, in the form of extrusion of a bulge, while the 5'-end and the T<sub>i</sub>C loop remain bound to the enzyme (Figure 4). Refolding thus leads to a structure that is reminiscent of a misaligned zipper with the bulged nucleotides protruding from an otherwise continuous RNA helix (Figures 2 and 3). This bulged substrate is stabilized by two new base pairs that form between the Cs added during the first CCA cycle and the 5'-terminal Gs and further stabilized by ionic interactions to the enzyme (Figure 3) as well as interactions between the bulge and loop 1 of the head domain. These interactions help lower the energy for the refolded conformation. (7–8) Once the active site is closed over the refolded RNA, the second cycle of CCA addition proceeds analogous to the first. Upon termination, the second CCA triplet again stacks against the head domain following a counter-clockwise rotation. (9) Nucleotide binding could then, in principle, result in yet another round of RNA refolding (10). However, RNA refolding at this stage would result in a 6nt bulge, a structure that is unable to accommodate a continuous helix, which is required for the RNA to remain bound to the enzyme (Figure S6). Therefore, the screw motion for active site closure and torque applied on the RNA following the second CCA-addition cycle induces dissociation of the RNA carrying CCACCA at its 3' end.

### Proofreading Unstable RNAs through Their Structural Versatility

Proofreading is an important feature of all polymerases. As DNA and RNA polymerases scan along a template strand while synthesizing a product strand, they recognize misincorporated nucleotides or erroneous templates either before or immediately after nucleotide incorporation (Kunkel and Bebenek, 2000; Sydow and Cramer, 2009). These transcriptional obstacles are usually removed through endonucleolytic cleavage before nucleotide synthesis resumes.

In contrast, the CCA-adding enzyme uses proofreading on two different levels. First, it is able to add the nucleotides CCA in a template-independent manner to the 3' ends of all tRNAs and tRNA-like molecules without translocating along its substrate (Pan et al., 2010; Toh et al., 2008; Tomita et al., 2006; Xiong and Steitz, 2004). Second, the enzyme exploits the structural versatility of unstable RNAs to mark them for degradation, even though the mutation or instability is never directly recognized by the active site of the enzyme. The enzyme simply sticks to its substrate requirements and catalytic principles regardless of the substrate. If a continuous A-form RNA helix of proper length is bound between the head and tail domains, and an unpaired discriminator base can be properly positioned in the active site, then the CCA-adding enzyme will add a CCA triplet to the RNA 3'-end. Rather than dissociating after the first CCA cycle, unstable RNAs are refolded in response to the regular screw motion of the enzyme during active site closure. CCA-addition terminates once the RNA substrate is unable to remain bound to the enzyme. We speculate that tandem CCA addition is the unavoidable consequence of CCA addition to a large subset of unstable RNAs beginning with GG, since refolding is triggered during catalysis. As it is driven by the RNA substrate itself and is beneficial for removing erroneous tRNAs, this mechanism is conserved across both classes of CCA-adding enzymes.

We suggest that RNA compression after completion of the first CCA cycle, which challenges RNA stability and triggers refolding of unstable RNAs, is an example of proof-reading via an energy relay mechanism, first proposed by Hopfield (1980). RNA compression as seen in the dynamic intermediate state of the enzyme is at a branch point, characteristic of these types of mechanisms. The RNA substrate at this branch point either “resists” and falls off the enzyme when it contains a stable acceptor stem, or “buckles,” extruding a bulge, and then receives another CCA when the acceptor stem is weakened due to a mismatch. The compression of the RNA provides the energy relay for proof-reading in this case.

Our study elucidates how non-coding RNAs (ncRNAs) that utilize the tRNA fold are either marked for rapid degradation or stabilized by the addition of a single CCA triplet. Since bona fide tRNAs possess very long half lives (Kanerva and Mäenpää, 1981), we speculate that CCA-addition is a way for ncRNAs to increase their stability. Interestingly, while the Men $\beta$  tRNA-like small RNA is rapidly degraded due to CCACCA addition in humans and mice, the Old World Monkey homolog is stable and thus a CCA target (Wilusz et al., 2011). In the future, it will be instructive to study the functional significance of these differences. Furthermore, the surveillance mechanism presented here is conserved across all kingdoms of life and across different classes of RNA. This underlines the centrality of RNA structure in regulating transcript stability and function (Ding et al., 2014; Rouskin et al., 2014; Wan et al., 2014) and calls for a more detailed search into the structural elements of lncRNAs.

## EXPERIMENTAL PROCEDURES

### Structures of the Archaeal CCA-Adding Enzyme in Complex with Unstable RNAs

The *A. fulgidus* CCA-adding enzyme (AfCCA) was expressed in *Escherichia coli* and all RNAs used for crystallography were purchased from Dharmacacon. Preinsertion complexes were obtained by co-crystallization with AMPcPP or CTP. Most reaction steps crystallized in 16%–24% PEG-3350 and 0.2–0.27 M sodium-potassium tartrate. Diffraction data were collected at beamline X25 of the National Synchrotron Light Source (NSLS) at Brookhaven National Laboratory. All structures were solved by molecular replacement using PHASER (McCoy et al., 2007) and refined with PHENIX (Adams et al., 2010). For details, see Extended Experimental Procedures.

### Structure of the Full-Length Human CCA-Adding Enzyme

The human mitochondrial CCA-adding enzyme was expressed in *E. coli* and crystallized in 2.3 M ammonium sulfate and 0.5 M tri-sodium citrate pH 5.8. Its structure was solved by molecular replacement and refined with PHENIX (Adams et al., 2010). For details, see Extended Experimental Procedures.

### Crosslinking Experiments, CCA-Addition Assays

A 2-amino-(S-trityl-ethylthio)-purine phosphoramidite was synthesized as described (Peacock et al., 2011). Crosslinkable RNA was prepared at ChemGenes. All in vitro assays were carried out with crosslinked and soluble components as described (Shi et al., 1998b). For details, see Extended Experimental Procedures.

### Competition Experiments

To study the effect of RNA pre-binding, full-length arginyl tRNA<sup>TCG</sup> (G70A) ending in its discriminator base was pre-incubated in the presence of cold CTP only. Cold ATP spiked with [ $\alpha^{32}$ P] ATP was added simultaneously with equimolar or excess amounts of full-length human Men $\beta$  already carrying

CCA. Reactions were allowed to proceed for 2 min for AfCCA and 30 s for the human CCA-adding enzyme. As a control experiment, both RNAs were added simultaneously along with ATP. For details, see [Extended Experimental Procedures](#).

#### Rosetta-Based RNA Modeling

*De novo* RNA folding was carried out using Rosetta 3.5 (<http://www.rosettacommons.org>) (Das et al., 2010). Detailed procedure in [Extended Experimental Procedures](#).

#### ACCESSION NUMBERS

Coordinates and structure factors of the archaeal CCA-adding enzyme complexes have been deposited in the Protein Data Bank (PDB) under accession codes 4X4N, 4X4O, 4X4P, 4X4Q, 4X4R, 4X4S, 4X4T, 4X4U, and 4X4V. The full-length human mitochondrial CCA-adding enzyme has been deposited under the accession code 4X4W.

#### SUPPLEMENTAL INFORMATION

Supplemental Information includes [Extended Experimental Procedures](#), six figures, four movies, and one table and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2015.01.005>.

#### ACKNOWLEDGMENTS

We thank Annie Héroux for help with data collection at beamline X25 at the National Synchrotron Light Source, which is supported by the Department of Energy, Office of Basic Energy Sciences, Andrei Laikhter from ChemGenes for RNA synthesis, Alan Weiner for the AfCCA expression plasmid, and Phillip Sharp and members of the L.J. laboratory for discussions and advice. We thank J. Jansen for comments on the manuscript. C.-D.K. was supported by a Postdoctoral Fellowship from the Jane Coffin Childs Memorial Fund for Medical Research. Early stages of this research were performed at MIT by J.E.W. supported in part by NIH grants R01-GM34277 and R01-CA133404 to P.S. This work was supported by NIH grants R00-GM104166 (to J.E.W.), R01-GM61115 (to P.A.B.), the Robertson Research Fund of Cold Spring Harbor Laboratory and the Cold Spring Harbor Laboratory Women in Science Award (to L.J.). L.J. is an investigator of the Howard Hughes Medical Institute.

Received: September 2, 2014

Revised: November 7, 2014

Accepted: December 31, 2014

Published: January 29, 2015

#### REFERENCES

Adams, P.D., Afonine, P.V., Bunkóczki, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* 66, 213–221.

Cho, H.D., and Weiner, A.M. (2004). A single catalytically active subunit in the multimeric *Sulfolobus shibatae* CCA-adding enzyme can carry out all three steps of CCA addition. *J. Biol. Chem.* 279, 40130–40136.

Cho, H.D., Verlinde, C.L., and Weiner, A.M. (2005). Archaeal CCA-adding enzymes: central role of a highly conserved beta-turn motif in RNA polymerization without translocation. *J. Biol. Chem.* 280, 9555–9566.

Clemson, C.M., Hutchinson, J.N., Sara, S.A., Ensminger, A.W., Fox, A.H., Chess, A., and Lawrence, J.B. (2009). An architectural role for a nuclear non-coding RNA: NEAT1 RNA is essential for the structure of paraspeckles. *Mol. Cell* 33, 717–726.

Das, R., and Baker, D. (2007). Automated *de novo* prediction of native-like RNA tertiary structures. *Proc. Natl. Acad. Sci. USA* 104, 14664–14669.

Das, R., Karanicolas, J., and Baker, D. (2010). Atomic accuracy in predicting and designing noncanonical RNA structure. *Nat. Methods* 7, 291–294.

Deutscher, M.P. (1982). tRNA nucleotidyltransferase. *Enzymes* 15, 183–215.

Ding, Y., Tang, Y., Kwok, C.K., Zhang, Y., Bevilacqua, P.C., and Assmann, S.M. (2014). In vivo genome-wide profiling of RNA secondary structure reveals novel regulatory features. *Nature* 505, 696–700.

Hopfield, J.J. (1980). The energy relay: a proofreading scheme based on dynamic cooperativity and lacking all characteristic symptoms of kinetic proofreading in DNA replication and protein synthesis. *Proc. Natl. Acad. Sci. USA* 77, 5248–5252.

Huang, H., Chopra, R., Verdine, G.L., and Harrison, S.C. (1998). Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. *Science* 282, 1669–1675.

Imamura, K., Imamachi, N., Akizuki, G., Kumakura, M., Kawaguchi, A., Nagata, K., Kato, A., Kawaguchi, Y., Sato, H., Yoneda, M., et al. (2014). Long noncoding RNA NEAT1-dependent SFPQ relocation from promoter region to paraspeckle mediates IL8 expression upon immune stimuli. *Mol. Cell* 53, 393–406.

Ji, P., Diederichs, S., Wang, W., Böing, S., Metzger, R., Schneider, P.M., Tidow, N., Brandt, B., Buerger, H., Bulk, E., et al. (2003). MALAT-1, a novel noncoding RNA, and thymosin beta4 predict metastasis and survival in early-stage non-small cell lung cancer. *Oncogene* 22, 8031–8041.

Kanerva, P.A., and Mäenpää, P.H. (1981). Codon-specific serine transfer ribonucleic acid degradation in avian liver during vitellogenin induction. *Acta Chem. Scand., B, Org. Chem. Biochem.* 35, 379–385.

Kunkel, T.A., and Bebenek, K. (2000). DNA replication fidelity. *Annu. Rev. Biochem.* 69, 497–529.

Li, F., Xiong, Y., Wang, J., Cho, H.D., Tomita, K., Weiner, A.M., and Steitz, T.A. (2002). Crystal structures of the *Bacillus stearothermophilus* CCA-adding enzyme and its complexes with ATP or CTP. *Cell* 111, 815–824.

Lin, R., Maeda, S., Liu, C., Karin, M., and Edgington, T.S. (2007). A large non-coding RNA is a marker for murine hepatocellular carcinomas and a spectrum of human carcinomas. *Oncogene* 26, 851–858.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. *J. Appl. Cryst.* 40, 658–674.

Nissen, P., Hansen, J., Ban, N., Moore, P.B., and Steitz, T.A. (2000). The structural basis of ribosome activity in peptide bond synthesis. *Science* 289, 920–930.

Pan, B., Xiong, Y., and Steitz, T.A. (2010). How the CCA-adding enzyme selects adenine over cytosine at position 76 of tRNA. *Science* 330, 937–940.

Peacock, H., Bachu, R., and Beal, P.A. (2011). Covalent stabilization of a small molecule-RNA complex. *Bioorg. Med. Chem. Lett.* 21, 5002–5005.

Rouskin, S., Zubradt, M., Washietl, S., Kellis, M., and Weissman, J.S. (2014). Genome-wide probing of RNA structure reveals active unfolding of mRNA structures in vivo. *Nature* 505, 701–705.

Shi, P.Y., Maizels, N., and Weiner, A.M. (1998a). CCA addition by tRNA nucleotidyltransferase: polymerization without translocation? *EMBO J.* 17, 3197–3206.

Shi, P.Y., Weiner, A.M., and Maizels, N. (1998b). A top-half tDNA minihelix is a good substrate for the eubacterial CCA-adding enzyme. *RNA* 4, 276–284.

Sprinzl, M., and Cramer, F. (1979). The -C-C-A end of tRNA and its role in protein biosynthesis. *Prog. Nucleic Acid Res. Mol. Biol.* 22, 1–69.

Sunwoo, H., Dinger, M.E., Wilusz, J.E., Amaral, P.P., Mattick, J.S., and Spector, D.L. (2009). MEN epsilon/beta nuclear-retained non-coding RNAs are up-regulated upon muscle differentiation and are essential components of paraspeckles. *Genome Res.* 19, 347–359.

Sydow, J.F., and Cramer, P. (2009). RNA polymerase fidelity and transcriptional proofreading. *Curr. Opin. Struct. Biol.* 19, 732–739.

Toh, Y., Numata, T., Watanabe, K., Takeshita, D., Nureki, O., and Tomita, K. (2008). Molecular basis for maintenance of fidelity during the CCA-adding reaction by a CCA-adding enzyme. *EMBO J.* 27, 1944–1952.

Tomita, K., Fukai, S., Ishitani, R., Ueda, T., Takeuchi, N., Vassylyev, D.G., and Nureki, O. (2004). Structural basis for template-independent RNA polymerization. *Nature* 430, 700–704.

Tomita, K., Ishitani, R., Fukai, S., and Nureki, O. (2006). Complete crystallographic analysis of the dynamics of CCA sequence addition. *Nature* **443**, 956–960.

Wan, Y., Qu, K., Zhang, Q.C., Flynn, R.A., Manor, O., Ouyang, Z., Zhang, J., Spitale, R.C., Snyder, M.P., Segal, E., and Chang, H.Y. (2014). Landscape and variation of RNA secondary structure across the human transcriptome. *Nature* **505**, 706–709.

Wilusz, J.E., Freier, S.M., and Spector, D.L. (2008). 3' end processing of a long nuclear-retained noncoding RNA yields a tRNA-like cytoplasmic RNA. *Cell* **135**, 919–932.

Wilusz, J.E., Whipple, J.M., Phizicky, E.M., and Sharp, P.A. (2011). tRNAs marked with CCACCA are targeted for degradation. *Science* **334**, 817–821.

Wilusz, J.E., JnBaptiste, C.K., Lu, L.Y., Kuhn, C.D., Joshua-Tor, L., and Sharp, P.A. (2012). A triple helix stabilizes the 3' ends of long noncoding RNAs that lack poly(A) tails. *Genes Dev.* **26**, 2392–2407.

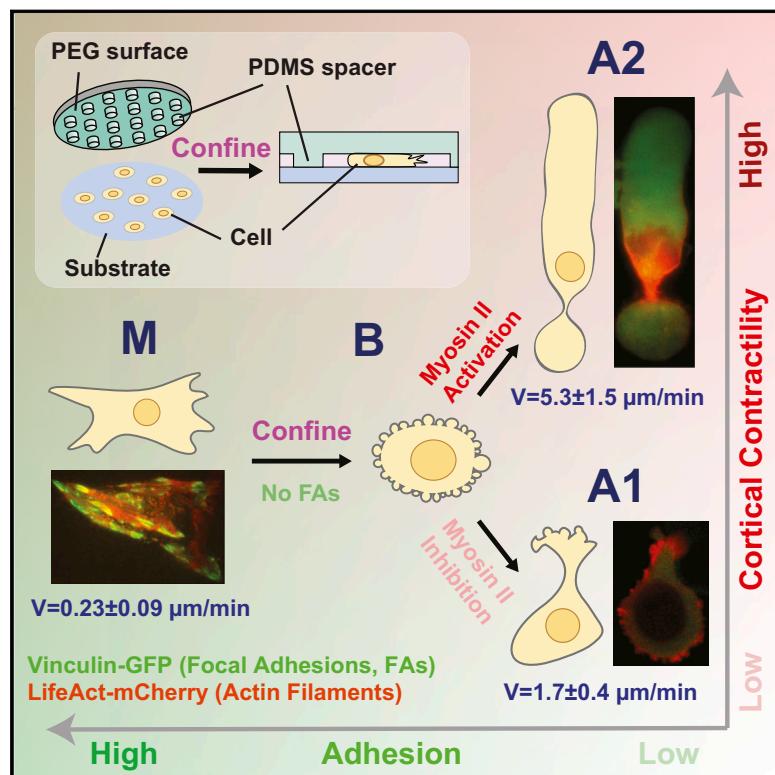
Xiong, Y., and Steitz, T.A. (2004). Mechanism of transfer RNA maturation by CCA-adding enzyme without using an oligonucleotide template. *Nature* **430**, 640–645.

Xiong, Y., Li, F., Wang, J., Weiner, A.M., and Steitz, T.A. (2003). Crystal structures of an archaeal class I CCA-adding enzyme and its nucleotide complexes. *Mol. Cell* **12**, 1165–1172.

Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31**, 3406–3415.

# Confinement and Low Adhesion Induce Fast Amoeboid Migration of Slow Mesenchymal Cells

## Graphical Abstract



## Authors

Yan-Jun Liu, Maël Le Berre, ..., Raphaël Voituriez, Matthieu Piel

## Correspondence

mael.leberre@curie.fr (M.L.B.), matthieu.piel@curie.fr (M.P.)

## In Brief

A large range of slow mesenchymal cells can switch to fast amoeboid-like migration under conditions of low adhesion and strong confinement, suggesting that tumor cells may spontaneously escape primary tumors and invade tissues without any specific genetic alteration.

## Highlights

- Physical confinement and low adhesion induce the mesenchymal-amoeboid transition
- A large range of slow mesenchymal cell types can display fast amoeboid-like migration
- A fast mode (A1) and a faster and more conserved contractile mode (A2) were observed
- A2 migration could be an ancestral migratory behavior shared among eukaryotes

# Confinement and Low Adhesion Induce Fast Amoeboid Migration of Slow Mesenchymal Cells

Yan-Jun Liu,<sup>1,6</sup> Maël Le Berre,<sup>1,6,\*</sup> Franziska Lautenschlaeger,<sup>1,3</sup> Paolo Maiuri,<sup>1</sup> Andrew Callan-Jones,<sup>5</sup> Mélina Heuzé,<sup>1</sup> Tohru Takaki,<sup>4</sup> Raphaël Voituriez,<sup>2</sup> and Matthieu Piel<sup>1,\*</sup>

<sup>1</sup>Institut Curie, CNRS UMR 144, 26 rue d'Ulm, 75005 Paris, France

<sup>2</sup>Laboratoire Jean Perrin and Laboratoire de Physique Théorique de la Matière Condensée, CNRS/Université Pierre et Marie Curie, 75005 Paris, France

<sup>3</sup>Universität des Saarlandes, Campus E2 6, 3. OG, Zi. 3.17, 66123 Saarbrücken, Germany

<sup>4</sup>Cancer Research UK London Research Institute, Clare Hall Laboratories, Blanche Lane, South Mimms, Hertfordshire EN6 3LD, UK

<sup>5</sup>Laboratoire Matière et Systèmes Complexes, CNRS/Université Paris Diderot, UMR 7057, 75204 Paris Cedex, France

<sup>6</sup>Co-first author

\*Correspondence: mael.leberre@curie.fr (M.L.B.), matthieu.piel@curie.fr (M.P.)

<http://dx.doi.org/10.1016/j.cell.2015.01.007>

## SUMMARY

The mesenchymal-amoeboid transition (MAT) was proposed as a mechanism for cancer cells to adapt their migration mode to their environment. While the molecular pathways involved in this transition are well documented, the role of the microenvironment in the MAT is still poorly understood. Here, we investigated how confinement and adhesion affect this transition. We report that, in the absence of focal adhesions and under conditions of confinement, mesenchymal cells can spontaneously switch to a fast amoeboid migration phenotype. We identified two main types of fast migration—one involving a local protrusion and a second involving a myosin-II-dependent mechanical instability of the cell cortex that leads to a global cortical flow. Interestingly, transformed cells are more prone to adopt this fast migration mode. Finally, we propose a generic model that explains migration transitions and predicts a phase diagram of migration phenotypes based on three main control parameters: confinement, adhesion, and contractility.

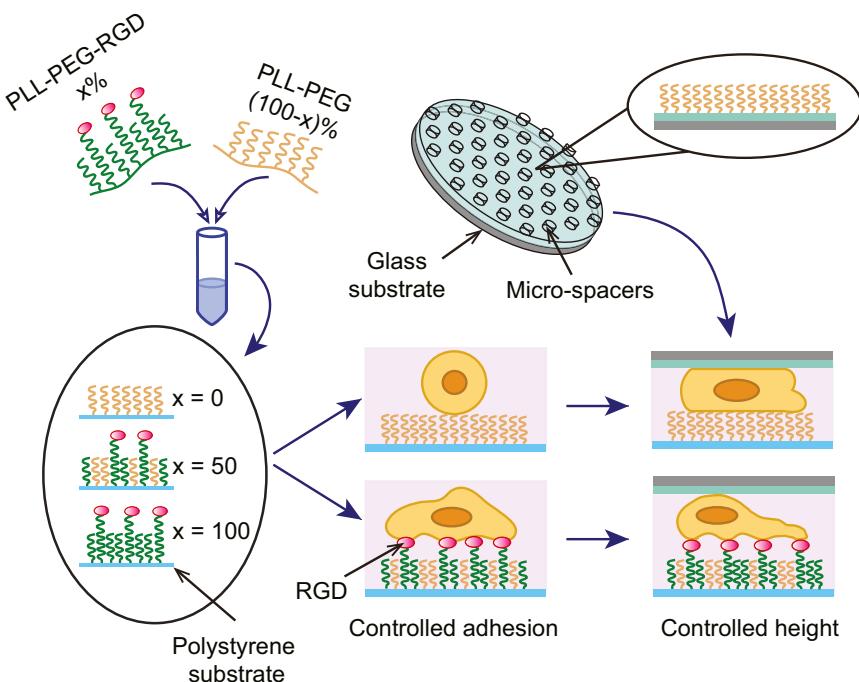
## INTRODUCTION

The mechanisms individual cells use to migrate can be divided into two distinct types, known as mesenchymal migration and amoeboid migration. Cells migrating in a mesenchymal fashion typically adopt an elongated, spindle-like shape and exert traction on their substrates via focal adhesions associated with actin rich protrusions, such as lamellipodia or filopodia. In contrast, cells undergoing amoeboid migration adopt round or irregular shapes. They undergo cycles of expansion and contraction, which allow them to squeeze through gaps in the extracellular matrix. The velocity of amoeboid migration is often higher than that of mesenchymal migration (Panková et al., 2010) and it appears to involve a range of different sub-modes, such as bleb-based migration or gliding (Paluch and Raz, 2013). Consequently, it is not as

well-defined at the molecular level as mesenchymal migration, although it is thought to rely on an increased activity of the RhoA pathway, which activates formin-based actin nucleation and myosin II contractility (Sanz-Moreno and Marshall, 2010).

Despite an extensive literature on the subject, it remains unclear whether the various modes of amoeboid migration are mechanistically distinct from each other, or whether they represent a continuum of a single process. Recent work, however, has identified mechanisms common to several forms of amoeboid migration, supporting the continuum model. Several studies report, for example, that amoeboid migration is often based on a contractile cell rear inducing a retrograde flow in the cell cortex (Heuzé et al., 2013; Paluch and Raz, 2013; Poincloux et al., 2011). Others have shown that, in contrast with mesenchymal migration, amoeboid migration does not depend on a cell's ability to form focal adhesions with its environment (Bergert et al., 2012; Lämmermann et al., 2008; Renkawitz et al., 2009). In the light of these studies, Lämmermann and Sixt (2009) have proposed that three key parameters modulate the mode of migration of cells, namely protrusion, contractility, and adhesion. They further suggest that by altering the balance between these parameters, cells could adapt their mode of motility to different migration environments (Lämmermann and Sixt, 2009), allowing migrating cells to negotiate the topologies of the different tissues they encounter.

Migration mode plasticity is an important factor for immune cells migration but also in cancer metastasis, as disseminating tumor cells have to navigate a range of extracellular matrix geometries to escape their point of origin and spread to distant sites. Metastasizing cancer cells exhibit both mesenchymal and amoeboid modes of migration, and recent studies have shown that they can switch between the two, a phenomenon known as the mesenchymal-amoeboid transition (MAT) (Friedl and Wolf, 2003; Panková et al., 2010; Taddei et al., 2013; Topluoglu et al., 2013). Importantly, non-migrating cells such as epithelial cells can first undergo a transition to mesenchymal migration (epithelial to mesenchymal transition or EMT), enabling them to develop collective and/or individual migratory behaviors associated with cancer cell dissemination (Friedl and Wolf, 2003; Taddei et al., 2013). This suggests that, by going through sequential transitions, even epithelial cells, could eventually display amoeboid motility (Giampieri et al., 2009).



**Figure 1. Schematic Diagram of Cell Migration Assay**

Cell culture substrates were treated with various ratios of pLL-PEG and pLL-PEG-RGD to control adhesion of cells with their environment and PEG-treated slides containing micro-spacers were used to confine cells at a fixed height. See also Figure S1.

density of RGD peptides on the cell substrate) (Barnhart et al., 2011) and to confine cells plated on a 2D substrate (Figure S1; Le Berre et al., 2012).

Normal human dermal fibroblast cells (NHDF) migrate in a mesenchymal fashion and so were chosen to study. We began by seeding NHDF cells on surfaces coated with various concentrations of pLL-PEG-RGD, but without confining them (Figure 2A). As RGD density decreased, cell spreading likewise decreased. Where RGD density dipped below 5%, cells were unable to attach to the substrate when gently flushed (Figures 2B and 2C).

To quantify differences in migration behavior, trajectories of individual cells (Figure 2D) were analyzed in terms of instantaneous speed (S), persistence (P) and diffusivity (D) (Figures 2E, S2D, and S2E; Movie S1; see Extended Experimental Procedures for details on calculation). In agreement with previous work by others (DiMilla et al., 1993), the long-term displacement of cells, which is characterized by their diffusion coefficient, is maximal for an intermediate level of adhesion (Figure 2F), due to an increase in persistence and a decrease in instantaneous speed with increasing adhesion (Figures 2G and 2H). These observations confirmed the typical mesenchymal migration of these fibroblasts.

#### Confinement Modifies the Migration Behavior of NHDF Cells

We introduced confinement by covering cells (cultured on surfaces of controlled adhesion) with a non-adhesive surface (Figure S1). NHDF cells were  $7.6 \pm 0.36 \mu\text{m}$  high during interphase. We therefore chose to limit our investigation to heights of either 5 or 3  $\mu\text{m}$ . At 5  $\mu\text{m}$  (low confinement), the cell body was slightly deformed. At 3  $\mu\text{m}$  (high confinement), both the cell body and the nucleus were deformed (Le Berre et al., 2012).

Cells tended to retract when confined, forming fewer large lamellipodial protrusions and more elongated pseudopodia (Movie S2; Figure 3A for 5  $\mu\text{m}$ , Figure S3A for 3  $\mu\text{m}$ , and phenotype C in Figures 3B–3E). This retraction of the cell margin was probably due to an increase in cell cortex contractility, which favors a more compact morphology (Sahai and Marshall, 2003). When adhesion was low ( $\leq 15\%$  RGD) and confinement high, cells became totally round, mostly immobile, and often showed a continuous blebbing activity (Movie S2; phenotype B in Figures

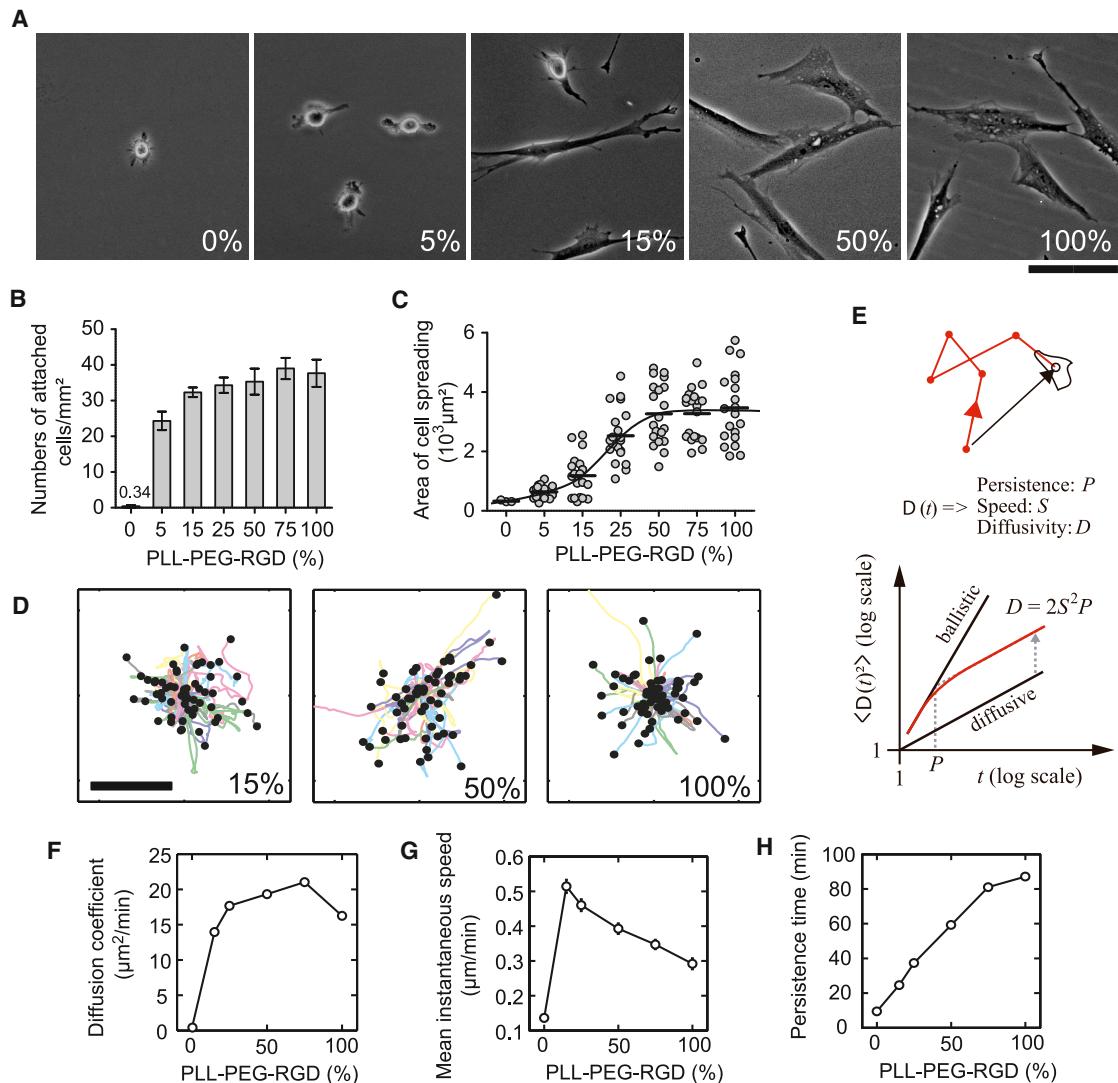
Given its importance in metastasis, the MAT has been the focus of several recent studies. These have identified key parameters driving the transition: cell-intrinsic factors, such as intracellular signaling and contractility, and extrinsic, environmental factors, such as cell adhesion and the geometry of the extracellular matrix. Using a protease inhibitor mix to inhibit the matrix remodeling ability of mesenchymal cells induces the MAT and is associated with a reduction in integrin expression (Carragher et al., 2006; Wolf et al., 2003). Tozluoglu et al. (2013) combined *in vivo* experiments and mathematical simulations to highlight cell confinement as a crucial parameter affecting the efficiency of various migration modes. *In vitro* studies using micro-channels also point to an important contribution of confinement in changing the requirements for efficient locomotion (Hung et al., 2013).

Until now, however, no systematic quantitative study has been performed to understand the interplay between a cell's environment and its migration behavior. Here, we describe how we systematically studied the effects of confinement and adhesion on migration modes. Using a variety of cell types, we found that slow mesenchymal cells can switch to fast amoeboid-like migration under conditions of low adhesion and high confinement. We propose a simple generic physical model that produces a phase diagram of migration modes and depends on both the intrinsic properties of cells and environmental parameters.

## RESULTS

### Adhesion Decreases Speed but Increases Persistence in Non-Confining Conditions

To systematically investigate the effects of adhesion and confinement on cell migration, we combined two methods (Figure 1) to control cell adhesion (by quantitative controlling of the



**Figure 2. Migration Behavior of NHDF Cells on Surfaces of Various Adhesive Levels**

(A) Representative phase contrast images of NHDF cells on surfaces containing various RGD densities.

(B and C) Quantification of the number of attached cells per mm<sup>2</sup> (B) and their spreading areas (C) for various RGD densities ( $n \geq 20$  cells,  $N \geq 2$  experiments for each point on the graph).

(D) Representative trajectories of 60 cells recorded over 24 hr for various RGD densities. The starting positions of each cell were registered to the center of the plot.

(E) Schematic explaining measurements of cell displacement  $D(t)$ , mean instantaneous speed ( $S$ ), and persistence ( $P$ ) of cells (see *Supplemental Information*).

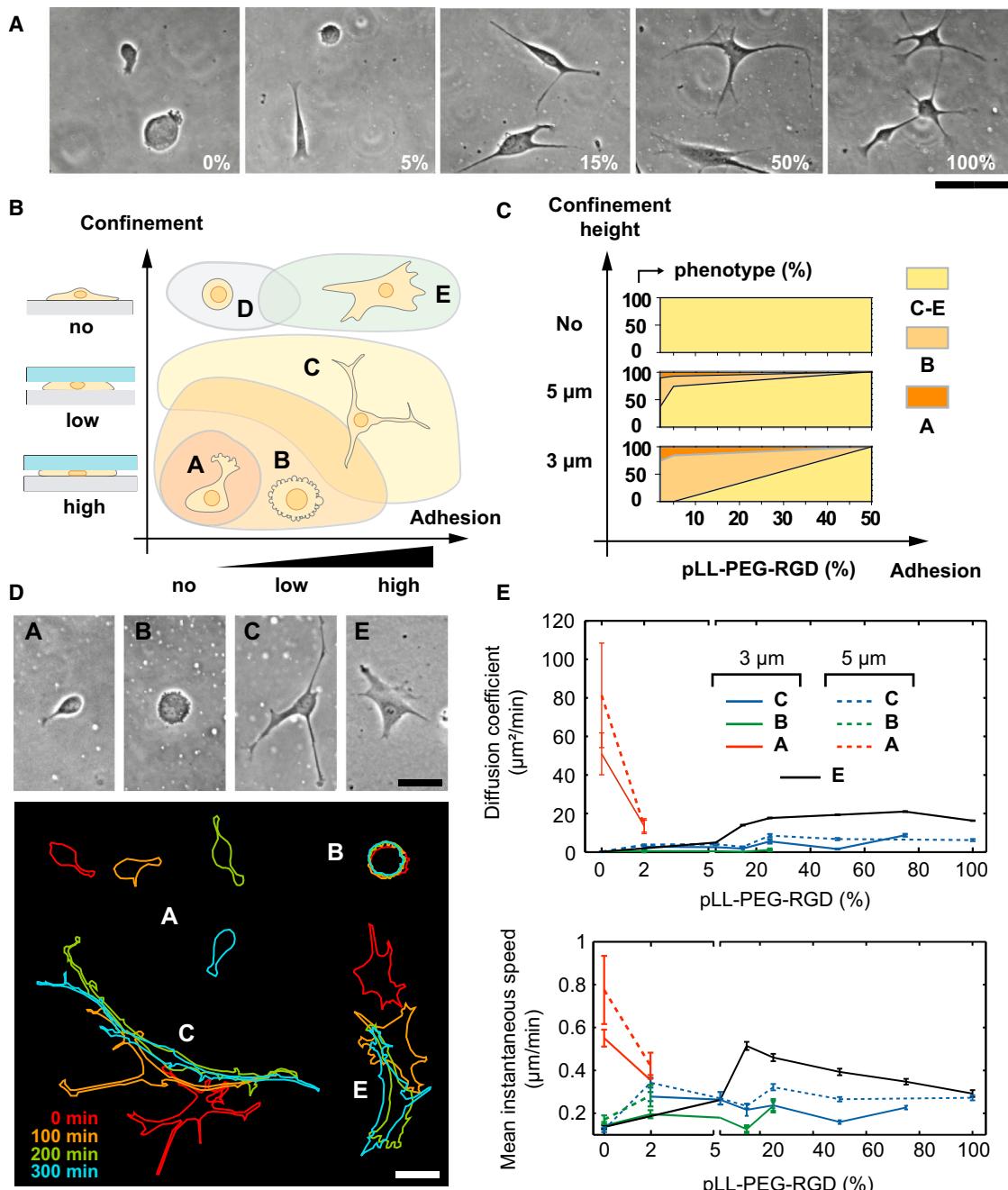
(F–H) Graphs showing diffusion coefficient (F), mean instantaneous speed (G), and persistence time (H) of NHDF cells for various RGD densities ( $n > 100$  cells for each point on the graph). In (A)–(H), % corresponds to the percentage of PLL-PEG-RGD in both PLL-PEG and PLL-PEG-RGD solution, error bars indicate SEM, and scale bars represent 100 μm.

See also Figure S2 and Movie S1.

3B–3E, note in Figure S3 legend), further suggesting an increased level of contractility. Unexpectedly, a fraction of these rounded NHDF cells could polarize and move fast with an amoeboid-like morphology (Movie S2; phenotype A in Figures 3B–3E).

Further quantification of the migration of spread cells (phenotype C or E) showed that confinement always reduced the instantaneous speed of such cells (Figure 3E), but did not significantly change their persistence (Figure S3B). Cells that were more rounded did not move, except for a sub-fraction of cells that spontaneously polarized (phenotype A). These cells showed

a much higher diffusion coefficient (up to ten times higher), which was not observed under conditions of strong adhesion or without confinement. The high diffusion coefficient of cells showing an amoeboid-like behavior could be attributed to both an increase in instantaneous speed and an increase in persistence. This is reminiscent of the behavior of leukocytes, such as dendritic cells, which are known to migrate fast and independently of specific integrin-based adhesion (Lämmermann et al., 2008; Lämmermann and Sixt, 2009; Hawkins et al., 2009; Heuzé et al., 2013).



**Figure 3. Migration Behavior of NHDF Cells in Response to Varying Confinement Height and Adhesion Strength**

(A) Representative phase contrast images of NHDF cells under 5  $\mu\text{m}$  confinement and various RGD densities. Scale bar represents 100  $\mu\text{m}$ .  
(B) Schematic showing the appearance of the observed phenotypes as a function of confinement height and adhesion strength. Phenotypes are defined as follows: A, round cell with a leading edge; B, round cell; C, cell with long and non-spread pseudopodia; D, floating, round cell; E, normally spread cell.  
(C) Portion of the different phenotypes observed under various conditions according to cell morphologies defined in (B).  $n \geq 100$  cells for each conditions.  
(D) Top panel: representative images of cells for each phenotype. Conditions of confinement ( $\mu\text{m}$ )/adhesion (% pLL-PEG-RGD) are A: 5/2%; B: 3/15%; C: 3/75%; E: no/100%. Bottom panel: corresponding time lapse of the cell outline. The color of the outlines represents elapsed time after the first outline. Scale bar represents 50  $\mu\text{m}$ .  
(E) Graph showing diffusion coefficient (top) and instantaneous speed (bottom) of each NHDF phenotype on varying RGD densities and for different confinements. Error bars indicate SEM.

See also Figure S3 and Movie S2.

Altogether, these observations show that a fraction of typically mesenchymal NHDF cells migrating under conditions of high confinement and low adhesion can switch to a fast amoeboid-like migration.

### **Two Distinct Modes of Amoeboid-like Migration Can Be Induced, Depending on Cell Type**

To test whether this microenvironment-dependent switch in the mode of migration was restricted to NHDF cells, we investigated the migration behavior of 20 other cell types confined by fully non-adhesive surfaces (100% pLL-PEG). We chose a sample of cell types covering a large range of migratory and non-migratory cells, including normal and transformed epithelial cells (pre- and post-EMT), normal and transformed mesenchymal cells, leukocytes (typical amoeboid cells), and muscle cell precursors.

For all cell types, with the exception of normal epithelial cells, we found a sub-fraction of spontaneously polarizing cells moving fast with an amoeboid-like morphology (Figures 4A and 4B). Unexpectedly, we found two distinct types of fast-migrating cells (Figures 4A and S4A–S4C). The first type, which we named A1, was similar to what was observed with NHDF cells: a round cell body and a small leading edge (Movie S3). The second, which we named A2, had an elongated ellipsoid cell body with a large uropod and resembled a migrating neutrophil (Figures S4A and S4B; Movie S3). Globally, A2 cells moved significantly faster ( $5.3 \pm 1.5 \mu\text{m}/\text{min}$ ) than A1 cells ( $1.7 \pm 0.4 \mu\text{m}/\text{min}$ ) and they both moved much faster than mesenchymal controls ( $0.234 \pm 0.09 \mu\text{m}/\text{min}$ ) (see Figures 4C and 4D).

The type of fast migration, A1 or A2, varied according to the cell line under observation (Figure 4D). For five cell types, the two modes could coexist in the same cell population, while most cell types preferentially switched to one migration mode or the other. Interestingly, different sub-clones of HeLa cells obtained from different labs (see *Supplemental Information* for a description of cell types) could either preferentially switch to A1 or A2 (Figure 4D). This suggests that these two migration modes, which are morphologically very different, can arise in very similar cell types, probably based on subtle molecular differences.

### **Transformed Cells, as well as Leukocytes, Preferentially Display Fast A2 Migration under Confinement**

Comparing different groups of cell types revealed general trends in the transition to amoeboid migration. Normal epithelial cells rarely displayed a transition to fast amoeboid migration (Figure 4D). This was due to the formation of large cell clusters that were mostly non-motile. At lower density, individual cells were able to display the A1 migratory phenotype.

When we induced EMT in MDCK or MCF10A cells with TGF $\beta$ , they no longer formed clusters and displayed fast A2 migratory phenotype. This was also true for tumor cell lines of epithelial origin from different tissues or in vitro transformed cell lines. Similarly, while normal mesenchymal cells tended to display mostly A1 migration under confinement, transformed mesenchymal cells often used in MAT studies displayed a high proportion of A2 migration mode when confined in a non-adhesive environment. Surprisingly, even muscle cell precursors displayed this fast migratory behavior under strong confinement. As expected, all leukocytes displayed A2-like migration,

including monocytes, which showed no motility on 2D substrates due to lack of adhesion. This suggested that the fast A2 migration mode was typically adopted by amoeboid cells, such as leukocytes, and by transformed cells migrating individually, independently of their tissue of origin.

Altogether, these results show that most slow mesenchymal cells can spontaneously switch to a fast amoeboid-like mode of migration when they are confined in a non-adhesive environment.

### **The Switch to Amoeboid-like Migration Depends on the Absence of Focal Adhesions**

The switch to amoeboid migration was observed only for adhesion levels that did not allow cell binding to the substrate when no confinement was applied (Figures 2B and 3C). To check for the presence of focal adhesions, we imaged HeLa VL cells by total internal reflection microscopy (TIRFM) (Trache and Meininger, 2008), on fibronectin-coated substrates and on PEG-treated substrates. On fibronectin, Vinculin-GFP displayed the expected pattern of focal adhesions, while on PEG, it was diffuse in the cytoplasm (Figure S5A). This was true both with and without confinement, showing that confinement did not restore the capacity of cells to form focal adhesions on PEG-treated surfaces.

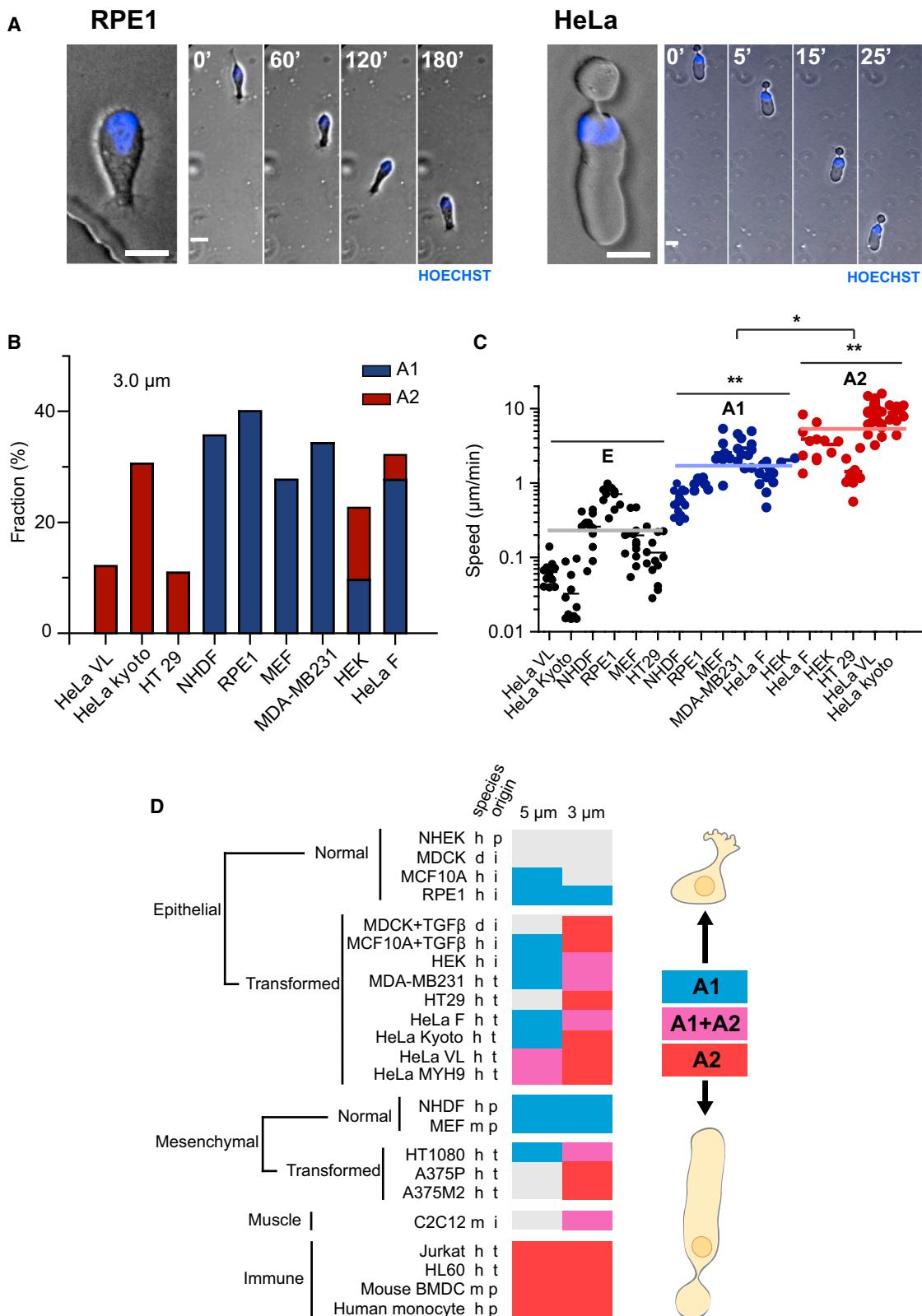
We then tested whether the ability of PEG-treated substrates to promote amoeboid migration relied on a lack of focal adhesions or on another surface effect of the PEG treatment. For this, we confined HeLa VL on an adhesive fibronectin surface and prevented formation of focal adhesions by knocking down the expression of Talin proteins with small interfering RNA (siRNA) or using the  $\beta 1$ -integrin-blocking antibody 4B4 (Figures 5, S5B, and S5F).

Under confinement on a fibronectin-coated surface, Talin knocked-down cells switched to A2 three times more often than cells treated with a non-targeting siRNA and as often as control cells on a non-adhesive PEG-treated substrate. Similarly, treating cells with 4B4 significantly increased the proportion of cells switching to amoeboid migration. This confirmed that the capacity to switch to amoeboid-like migration depended on the absence of focal adhesions and not on the chemical nature of the migration substrate or on a particular molecular perturbation of focal adhesion complexes.

### **The Switch to Fast A2 Amoeboid-like Migration Depends on High Cell Contractility, whereas A1 Requires Low Contractility**

The comparison of a large number of cell types under two different levels of confinement suggested that increasing confinement led to a larger proportion of A2 migrating cells (Figure 4D). We confirmed this effect using three different levels of confinement for HeLa VL cells. While most migrating cells displayed A1 migration under low confinement, only A2 was observed under high confinement (Figure 5B).

To investigate the role of the myosin II motor protein in the behavior of confined cells, we first treated confined HeLa VL cells with increasing doses of the phosphatase inhibitor calyculin A. This inhibits the dephosphorylation of myosin II, thereby increasing its activity (Ishihara et al., 1989). Increasing calyculin A concentration increased the fraction of cells switching to A2



(legend on next page)

migration mode, up to 66% of the total population (Figures 5C and 5D; Movie S4). We also used siRNA to knock down the expression of MYPT1 (Figure S5F), the PP1 partner targeting myosin II (Feng et al., 1999). MYPT1 siRNA induced an increase in the volume and number of blebs (Figure S5C) and an increase in the fraction of cells switching to A2 (Figures 5E and S5D), confirming the role of myosin II-based contractility in this phenomenon.

High-resolution imaging showed that the transition to A2 corresponded to the formation of a large bleb where most of the cell cytoplasm, including the nucleus, was transferred (Figure 5F; Movie S5), inducing a profound reorganization of the cytoskeleton and resulting in a new stable shape with an elliptical cell body and a rear uropod. The formation of such large blebs occurred at a constant rate (Figure S5E) and more frequently when cells were treated with calyculin A or knocked-down for MYPT1 (Figure S5E). Overall, this suggests that confinement favors cell contractility and that high cell contractility favors fast A2 migration.

To confirm that hypothesis, we measured the contractility of cell populations displaying different fractions of A2 phenotype (see **Supplemental Information**). We plated cells between two non-adhesive gels of known rigidity (1 kPa) coated with fluorescent beads (Figure S5H). Using 3D high-resolution imaging, we assessed the extent to which the cells deformed the gels and so deduced the ability of cells to push on their surroundings, as a measure of their relative level of cortical contractility (Figure S5I). This measure increased with increasing myosin II activity (Figures 5H and S5G), and stronger cells displayed a higher fraction of A2 migratory phenotype (Figure 5H). We then assayed other cell types and found that cells displaying A2 migration were significantly stronger than cells displaying A1 migration (Figure 5I). This confirmed our hypothesis that A2 migration appeared above a certain threshold of contractility.

Accordingly, treating cells with the ROCK inhibitor Y27632 reduced bleb formation and the proportion of cells switching to A2. More generally treatment with blebbistatin prevented formation of A2 migrating cells in all cell types tested. Unexpectedly, however, it did not prevent HeLa cells from moving and instead drastically increased the number of cells switching to the A1 amoeboid-like migration mode (Figures 5C and 5G; Movie S4).

Together, these experiments show that high contractility favored fast A2 migration mode, by increasing blebbing activity, eventually leading to the formation of large “stable blebs” comprising most of the cell content. Conversely, low contractility favored the formation of protrusive structures, leading to A1 migration.

## A2 Migration Mode Relies on a Fast and Global Cortical Retrograde Flow, whereas A1 Migration Relies on Local Actin Polymerization

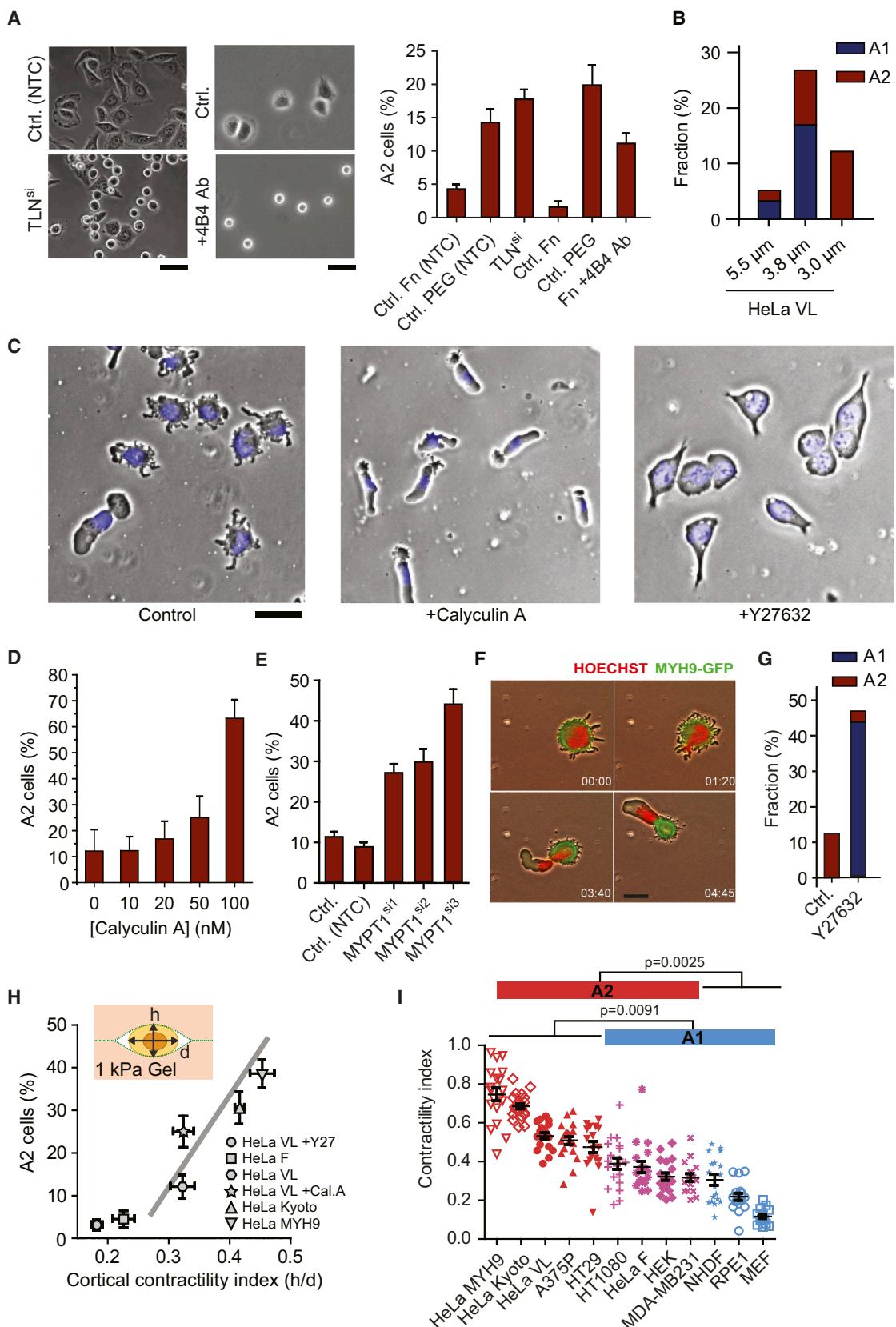
To characterize the migration mechanism of confined non-adhesive cells, we imaged focal adhesions, actin, and myosin II using TIRFM. On the fibronectin-coated surface, cells attached, spread, and formed focal adhesions and actin stress fibers containing myosin II (Figure 6A, upper panel; Movie S6). In A2 cells migrating under confinement on a non-adhesive surface, Vinculin-GFP was homogeneously distributed in the cytoplasm, and no stress fibers could be observed. Instead, actin filaments assembled into a fibrous cortex that was absent from the cell front, but became progressively denser toward the cell rear at the uropod (Figure 6A, middle panel; Movie S6). Myosin II formed patches in A2 cells, mostly on the rear part of the cell where the actin cortex was denser. Both actin filaments and myosin II patches showed a strong retrograde flow in the central part of the cell ( $15.8 \pm 3.7 \mu\text{m}/\text{min}$  SD n = 11, measured in the reference frame of the cell), whereas almost no actin movement was observed in the rear of the cell and in the uropod (Figure 6A bottom panel; see also Movie S7 and Figure 6B for flow visualization by particle image velocimetry).

To assess whether A1 motility relied on a similar mechanism, we added 30  $\mu\text{M}$  of the ROCK inhibitor Y27632. As with the A2 cells, we found that motile single A1 cells displayed a fast retrograde flow ( $14.63 \pm 0.9 \mu\text{m}/\text{min}$ ). Unlike the A2 cells, however, the flow was not global, from the front to the back of the cell. Instead, it was localized only in a small protruding region at the leading edge of the cell (Figures S6A and S6B; Movie S8). The limited area of the retrograde flow might explain why these cells were slower than A2 cells.

Several studies have shown that such a retrograde flow is responsible for cell movement independently of adhesion in amoeboid cells (Berger et al., 2012; Renkawitz et al., 2009). To confirm the role of the retrograde flow in cell movement, we also compared retrograde flow speed with cell speed and found a significant correlation between these two parameters (Figure S6E). However, retrograde flow was faster than the cell speed, indicating that a limited amount of friction occurs between the flowing cortex and the substrate surface (Renkawitz et al., 2009). In A2 cells lacking a uropod, the cell could go as fast as the retrograde flow (Figure S6F; Movie S8), and on average, A2 cells were 1.84 time faster (1.32–2.58 at 95% confidence) after being detached from their uropod (see also Figures S6C and S6D). These results confirmed that A2 cells moved via a coupling of a fast acto-myosin retrograde flow with the substrate and that the uropod acted as a dragged passive body.

**Figure 4. Strong Confinement in a Non-Adhesive Environment Can Induce Two Distinct Modes of Amoeboid Migration**

- (A) Time-lapse sequences of a typical amoeboid RPE1 (left) and HeLa VL (right) cell under 3- $\mu\text{m}$  confinement on a non-adhesive PEG surface. Grey is phase contrast and blue is nuclear staining (Hoechst). Time is in min ('). Scale bar represents 20  $\mu\text{m}$ .
- (B) Fraction of each phenotype for various cell lines under 3- $\mu\text{m}$  confinement on a non-adhesive surface (n = 140, 147, 267, 129, 60, 138, 137, 115, and 112 cells).
- (C) Instantaneous migration speed for various cell lines depending on their migratory phenotype (n = 13, 13, 13, 13, 13, 13, 22, 6, 10, 11, 10, 2, 9, 3, 10, 46, and 11). \*p < 0.05, \*\*p < 0.01 according to t test.
- (D) Comprehensive table indicating the presence or absence of amoeboid phenotypes for various cell lines: blue indicates RPE1-like (A1), red indicates HeLa-like (A2), purple indicates both (A1+A2), and gray indicates none. Cells are sorted by their origins and state of transformation. Species are human (h), mouse (m), and dog (d). Origins are primary (p), immortalized (i), and tumoral (t). See also Figure S4 and Movie S3.



(legend on next page)

## In A2 Cells, the Stable and Global Retrograde Flow Is Induced by Myosin II Contractility and Asymmetrical Actin Turnover

To better understand how global retrograde flow could be maintained and how actin was recycled in the cell, we calculated the polymerization dynamics of filamentous actin. We first measured the 1D retrograde flow of actin and myosin II along the length of cells (see [Extended Experimental Procedures](#) for details). In individual cells, the mean rate of flow (in the reference frame of the moving cell) was similar for both proteins ([Figure 6C](#)). In terms of protein localization, however, myosin II patches appeared to lie further toward the back of the cell ([Figure 6D](#); [Movie S6](#)). Interestingly, we noticed that myosin II was localized where the actin retrograde flow accelerated ([Figure 6D](#), bottom-right graph) and that the gradient of speed of the actin filaments, which accounts for the actin network contraction, was similar to the mean myosin II density along the cell ([Figure 6D](#)). This suggested a role for myosin II in accelerating the retrograde actin flow.

To produce a stable cortical flow, cortical actin has to be recycled. According to the well-accepted model for actin filament turnover ([Wilson et al., 2010](#)), new actin monomers are polymerized at the cell front and filamentous actin is severed at the back of the cell. As in [Wilson et al. \(2010\)](#), we were able to use the mass conservation law to deduce the rate of polymerization of filamentous actin along the cell length (see [Supplemental Information](#) for details). We found that actin polymerized at the front at a constant rate and depolymerized at the back ([Figure 6E](#)). A similar analysis of myosin II dynamics showed that recruitment of this protein was reduced toward the cell front, suggesting that it was rapidly captured by actin filaments and moved toward the rear of the cell.

Based on these elements, a mathematical model can be proposed (summarized in [Figure 6H](#) and [Supplemental Information](#)) (see also [Ruprecht et al. \[2015\]](#) in this issue of *Cell*; [Callan-Jones and Voituriez, 2013](#); [Hawkins et al., 2011](#)), which predicts

appearance of a motile “stable bleb” state above a threshold of contractility of the acto-myosin cortex.

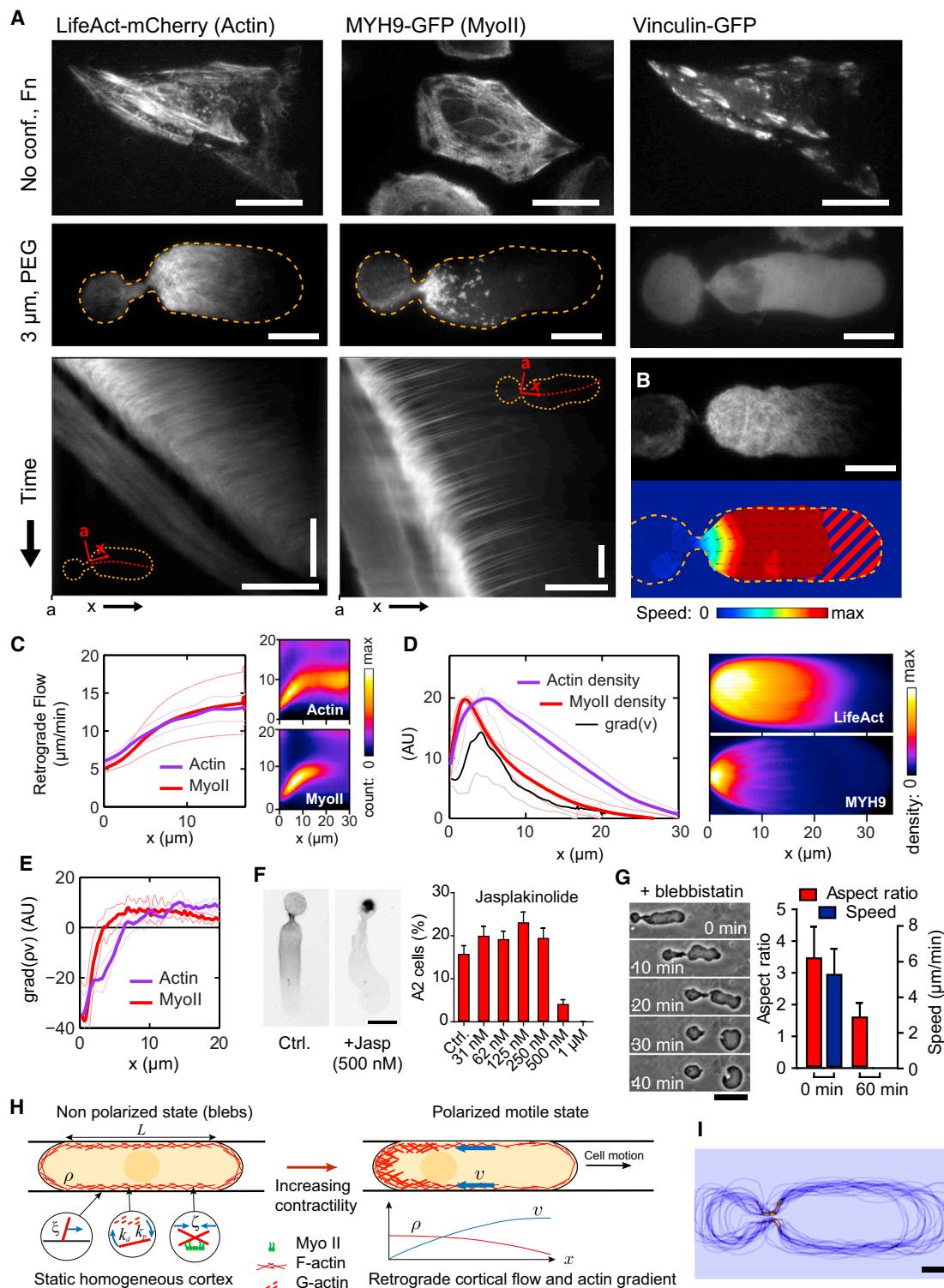
To further test this model, we assessed the respective roles of actin turnover and myosin II contractility in A2 migrating cells. When cells were treated with low concentrations of jasplakinolide, a drug that interferes with actin filament turnover, an increasing fraction of the actin filaments was found to accumulate at the cell back ([Figure 6F](#)). Above 250 nM, the proportion of A2 cells dropped ([Figure 6F](#)). Treated cells also displayed an unstable and floppy front, showing that cortex turnover was essential for the stability of the A2 mode. To test the role of myosin II during steady-state migration, we introduced the drug into the device after the A2 cells had formed. After treatment, cells concomitantly lost their elongated shape to become round and stopped moving ([Figure 6G](#)). Myosin II activity was thus not only required for spontaneous polarization, but also, together with actin filament turnover, to maintain the cortical flow, the shape, and the motility of the A2 mode.

All these observations are consistent with our model and are identical to the observations made on embryonic fish cells reported by [Ruprecht et al. \(2015\)](#) in this issue of *Cell*. This suggests a common generic mechanism for the migration of contractile confined cells which are not able to form actomyosin stress fibers and instead form a contractile cortex. When a cell exceeds a certain level of contractile activity, its cortex can be destabilized by blebs and switch spontaneously to another stable state that induces a fast migration behavior ([Figure 6I](#)). Conversely, when a cell is only moderately contractile, its cortex is constantly being destabilized by blebs but does not switch to a new stable configuration, leading to non-migrating, blebbing cells. When contractility is completely inhibited, the cortex is not destabilized, and a cell can develop more protrusive activity, which eventually leads to the A1 type amoeboid-like migration.

**Figure 5. Migratory Transitions Depend on Myosin II Activity and on the Presence of Focal Adhesions**

- (A) Left panel: representative images of HeLa VL on a fibronectin-treated substrate after treatments against adhesion: by Talins depletion (TLN1+TLN2) with siRNA compared with a non-targeting control (NTC) or by blocking  $\beta 1$  integrins with the 4B4 antibody (30  $\mu$ g/ml). Right panel: fraction of A2 cells after confinement on a non-adhesive surface (PEG) or a fibronectin-coated surface (Fn) with or without treatments against focal adhesions ( $n = 482, 313, 687, 240, 424$ , and  $436$  cells). Scale bar represents 50  $\mu$ m. All confined cells are under 3- $\mu$ m confinement and error bars are either SEM for means or error margin for %.
- (B) Fraction of each phenotype for HeLa VL cells depending on confinement level on a non-adhesive PEG surface ( $n = 308, 254$ , and  $140$  cells). All confined cells are under 3- $\mu$ m confinement and error bars are either SEM for means or error margin for %.
- (C) Representative images of HeLa VL cells without and with 100 nM calyculin A or 30  $\mu$ M Y27632 treatments under 3- $\mu$ m confinement on a non-adhesive PEG surface. Scale bar represents 50  $\mu$ m. All confined cells are under 3- $\mu$ m confinement and error bars are either SEM for means or error margin for %.
- (D) Fraction of A2 cells for various calyculin A concentrations, in HeLa VL cells ( $n = 140, 327, 208, 140$ , and  $185$  cells). All confined cells are under 3- $\mu$ m confinement and error bars are either SEM for means or error margin for %.
- (E) Fraction of HeLa VL cells transitioning to the A2 phenotype after MYPT1 depletion with three distinct sequences of siRNAs or after a non-targeting control (NTC) treatment ( $n = 664, 637, 408, 211$ , and  $177$  cells). All confined cells are under 3- $\mu$ m confinement and error bars are either SEM for means or error margin for %.
- (F) Time-lapse sequences of a HeLa Kyoto cell showing transition from a rounded blebbing cell (phenotype B) to A2 phenotype. Images are an overlay of phase contrast (gray), MYH9-GFP fluorescence (green, staining myosin II), and Hoechst nuclear staining (red). Time is in hr:min. Scale bar represents 25  $\mu$ m.
- (G) Fraction of HeLa VL cells transitioning to A1 and A2 phenotypes after 30  $\mu$ M Y27632 treatment ( $n = 140$  and  $191$  cells). All confined cells are under 3- $\mu$ m confinement and error bars are either SEM for means or error margin for %.
- (H) For various HeLa clones or after treatment with Y27632 (30  $\mu$ M) or calyculin A (50 nM), fraction of A2 cells compared with the mean contractility index of cells <math>\langle h/d \rangle</math>, which quantifies the ability of detached cells to deform a soft non-adhesive polyacrylamide (PA) gel of 1 kPa. For each point:  $n > 200$  cells on y axis and  $n > 15$  on x axis.
- (I) Contractility index of various cell types showing either A1 or A2 migration ( $n = 19, 22, 18, 20, 17, 19, 18, 19, 17, 18, 17$ , and  $15$ ; p values are from a non-parametric Mann-Whitney test).

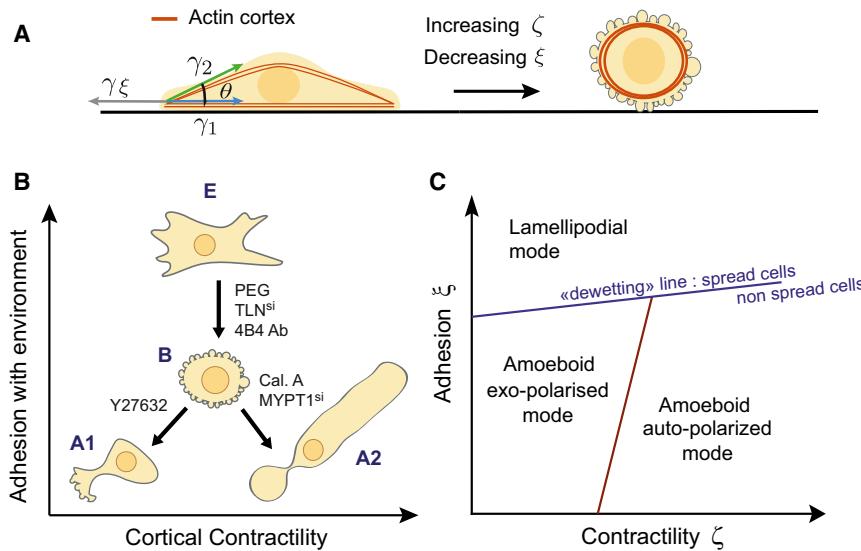
See also [Figure S5](#) and [Movies S4](#) and [S5](#).



**Figure 6. Motility of A2 Fast-Moving Phenotype Is Based on a Strong Actin Retrograde Flow**

(A) Representative TIRF microscopy images of HeLa cells expressing both Vinculin-GFP (left) and LifeAct-mCherry (right) (HeLa VL), and HeLa cells expressing MYH9-GFP (middle) under the following two conditions: on fibronectin-coated adhesive surface without confinement (top panel) and confined under 3 μm on a non-adhesive PEG surface (middle panel). Bottom images correspond to kymographs showing retrograde flow in the cell body. Orange dashed lines show cell outlines, red dashed lines were used to generate kymographs. Horizontal bars are 10 μm and vertical bars are 2 min.

(legend continued on next page)



**Figure 7. Schematic of the Proposed Model Describing Transitions between Modes of Cell Migration**

(A) Increasing cortical tension and decreasing cortex-surface interaction both converge toward a round, “non-wetting” shape of the cell.

(B) Summary of experimental observations.

(C) Generic phase diagram resulting from models of cortical wetting and cortical stability.

## DISCUSSION

### A Phase Diagram of Cell Migration Can Be Drawn Based on Only Three Parameters

Amoeboid migration has been described as a fast migration mode potentially independent of adhesion and involving a rounded cell shape. However, no consensus on its mechanism has been established. Here, we showed that physical confinement and low adhesiveness could promote a mesenchymal-to amoeboid transition across a range of cell types that have not been thought to be capable of amoeboid motility. In addition, cortical contractility determined the type of amoeboid migration, leading to a general phase diagram for single cell locomotion, based on only three main parameters: adhesion, confinement, and contractility (Figure 7B).

Below, we discuss in more detail how these three parameters, either individually or in conjunction, might contribute to the migratory transitions that we observed.

### The Absence of Focal Adhesion Allows the Formation of a Cell Cortex and a Switch to a Rounded State

An intriguing aspect of the transition to amoeboid migration modes is that, in all cases, it required the absence of focal ad-

hesions. We found that when cells were still able to form focal adhesions, they formed persistent protrusions and stress fibers, all of which prevented the formation of large-scale cortical flows. This observation is consistent with previous reports that integrin expression inhibits amoeboid migration (Carragher et al., 2006). Focal adhesions and the cell cortex might compete to recruit the acto-myosin contractile machinery, leading to either contractile stress fibers or a contractile/flowing cortex. In the absence of focal adhesions, cortical tension can trigger the mechanical delamination of remaining adhesive sites, leading to a rounded shape. To mathematically account for this transition from a spread, adherent phenotype to a round shape at low adhesion and high contractility, we can draw an analogy with the dewetting of liquid droplets on flat surfaces (see *Supplemental Information*). This qualitatively accounts for the observed transition from phenotype E to B in Figure 7. This analogy, although speculative at that stage, is useful to complete the theoretical framework recapitulating our observations. It captures the importance of the balance between contractility and adhesion.

Importantly, adhesion alone does not prevent the transition to amoeboid migration, but rather its association with the formation of stress fibers. As a consequence, cell types that could adhere to particular matrix components without generating stress fibers could also move with a flowing cortex, leading to amoeboid migration even in the presence of a strong coupling between the actin cortex and the cell environment. This type of motility has been observed in leukocytes, in cancer cells invading a soft non-crosslinked matrix such as Matrigel (Poincloux et al.,

- (B) Particle image velocity (PIV) of actin flows (bottom image) recorded on a TIRF movie (top image). Dashed part corresponds to areas where signal was too weak for flow calculation. Scale bar represents 10  $\mu\text{m}$ .
- (C) Mean retrograde flow along the cell (in the reference frame of the cell). Right images show speed distribution measured for 15 min on an individual cell and left graph shows mean flow averaged over several cells.  $x = 0$  corresponds to the position of the neck.
- (D) Mean actin (based on LifeAct staining) and myosin II density profiles. Right images show 2D mean densities in the cell body averaged for 15 min on a single cell and left graph shows mean density profile for several cells. Myosin II density colocalized with actin retrograde flow gradient grad(v).
- (E) Polymerization rates estimates along the cell body for myosin II and actin. Inferred from density and retrograde flow profiles shown in (C) and (D).
- (F) Left: actin localization (inverted image of LifeAct fluorescence) in a HeLa VL A2 cell with or without jasplakinolide treatment. Right: fraction of A2 cell function of jasplakinolide concentration (for each bar  $n > 278$  cells, error bars are error margins). Scale bar represents 20  $\mu\text{m}$ .
- (G) Time-lapse sequence of an A2 HeLa VL cell after a 150  $\mu\text{M}$  blebbistatin treatment (left) and quantification of the aspect ratio of the cell front and cell speed before and 60 min after treatment ( $n = 7$  cells, error bars are SEM) (right). 3- $\mu\text{m}$  confinement, PEG. Scale bar represents 50  $\mu\text{m}$ .
- (H) Schematic of the proposed model for cell motion (see main text and *Supplemental Information*).
- (I) Mean cell shape: superposition of the outlines of 22 cells showing the reproducibility and stability of cell shape in mode A2. Scale bar represents 10  $\mu\text{m}$ . In (C)–(E),  $n = 5$  (actin), six (myosin II) cells, error envelope is SEM and 0 on x axis corresponds to the neck position. Point a in (A).

See also Figure S6 and Movies S6, S7, and S8.

2011), and in embryonic cells from zebrafish (Ruprecht et al. [2015] in this issue of *Cell*).

### Confinement Is Required for Transmission of Forces in the Absence of Adhesion, but It Also Increases Contractility

When mesenchymal cells are prevented from forming focal adhesions, they cannot hold on to fibronectin- or collagen-coated substrates. This automatically makes confinement a requirement for cell migration. This phenomenon has been termed “chimneying” (Malawista et al., 2000), in reference to a technique used by mountain climbers. Exactly how these frictional forces are transmitted from the acto-myosin cortex to the migration substratum remains an open question. In addition, it has been proposed that confinement, in more complex geometries, combined with blebbing activity and an intrinsic cell polarity, could allow complete non-frictional motion to occur (Tozluoglu et al., 2013). In conclusion, the most direct function of confinement is to enable force transmission in the absence of adhesions to the substrate.

An unexpected effect of confinement was to significantly increase contractility, thus favoring blebbing and the A2 mode of migration. As this effect was immediate, it could not rely on the activation of a transcriptional program. Force-induced signaling via focal adhesions could not explain this effect either, as cells did not form adhesive structures. This suggests that the increase in contractility observed upon confinement might be induced directly by cell and/or nuclear deformation through a yet-unknown mechanism.

### A Major Difference between A1 and A2 Migration Modes Is the Origin of the Polarization Process that Triggers Cell Movement

An intrinsic polarization factor or process in the cell, or an extrinsic spatial cue can induce a local retrograde flow in a protrusive structure (Lorentzen et al., 2011). Indeed, we observed that cells switched to the A1 migration mode after making contact with cell debris or with other cells, or after a division. This indicated that this mode of migration only rarely appeared spontaneously and relied on a polarization factor to locally trigger the retrograde flow and thus the cell movement.

Cell polarity can also be induced in the absence of a specific polarization signal, a phenomenon previously described in other studies (Lorentzen et al., 2011; Poincloux et al., 2011). We found that this spontaneous polarization of a contractile cell cortex, which is predicted by our minimal model based on active gel theory (see *Supplemental Information*), can be nucleated randomly by the blebbing activity (see also Ruprecht et al. [2015] in this issue of *Cell*). Following destabilization of the cortex and transfer of the cytoplasm to the large bleb, contractility remains concentrated at the back of the cell due to the global flow, generating a stabilizing feedback. This can trigger an increase in pressure in the cell, a phenomenon reminiscent of lobopodial migration (Petrie et al., 2012, 2014). The movement of the global actin retrograde flow can also be transmitted to the substratum.

This phenomenon of a steady-state flow based on actin filament turnover is essentially similar to the mechanism driving pro-

trusive lamellipodial migration (Wilson et al., 2013) or A1 cell motility. But, importantly, while A1 motility requires an additional mechanism for polarization, in the case of high contractility (A2 mode), we provided evidence, both theoretical and experimental, that a stable polarization and thus a persistent migration can occur spontaneously.

### A2 Migration: A Generic, Ancestral Mode of Cell Locomotion?

For all cell lines studied, the speed of amoeboid-like migration was much faster than the speed of the mesenchymal form and corresponded to values reported for amoeboid cell types (several  $\mu\text{m}/\text{min}$ ). This suggests that these typical speeds are rather linked to the mode of migration than to the type of cell, as a given cell type can adopt both types of migration and associated speed, depending on its migration environment.

While we could observe two different types of fast amoeboid migration, the A2 mode appeared to be more general. Most cell types were able to adopt it, if contractile and sufficiently confined. Strikingly, all cell lines displayed a unique shape and speed when in the A2 mode, resembling the shape of a migrating neutrophil, with a pronounced uropod and a smooth rounded leading edge and with a similar range of speed. This was independent of their tissue of origin, of their physiological function, and of the shape and speed they displayed when migrating on an adhesive substrate without confinement. Importantly, embryonic cells from zebrafish could also display the same general shape and range of speed (Ruprecht et al. [2015] in this issue of *Cell*).

When we investigated the acto-myosin dynamics underlying the migration of these cells in more detail, the similarity appeared to be even more striking. Interestingly, the stereotypical “sausage” shape of the A2/stable bleb migrating cells has also been reported for amoebae, with also a similar actin and myosin II distribution (Yoshida and Inouye, 2001). This suggests that the transition to the A2, or “stable-bleb” fast amoeboid-like migration might be a very general phenomenon, relevant not only for cultured cancer cells but also for primary embryonic cells, both *in vitro* and *in vivo*.

From a mathematical point of view, the transition to A2 migration mode can be understood with a very generic minimal 1D theoretical model. The fact that such a simple and generic model can predict the transition to A2/stable bleb migration suggests that this mode of locomotion could be viewed as a basic property of a dynamic and contractile actin cortex. Consistent with this view, we often observed cell fragments of various sizes, resulting from shedding of large blebs, moving with the same shape, speed, and actin dynamics as A2 cells.

Taken together, the generality of the A2/stable bleb migration mode, the strong conservation of its main features (cell shape, cell speed, acto-myosin distribution, and dynamics) across a variety of cell types and organisms ranging from amoebae to vertebrates, the generic aspect of the model describing it, and the key feature of spontaneous polarization, make it tempting to speculate that this mode of migration may be an ancestral locomotion mechanism for crawling cells, shared by eukaryotes.

## A2 Migration Mode Is Related to Cell Transformation

An interesting result emerging from the comparison of various cell types (Figure 4D), is that the capacity to display amoeboid motility seems to correlate with the transformed character of the cells assayed. This suggests an increased migration plasticity of cancer cells.

Importantly, the various migration modes can be observed in a single cell line by slightly varying the confinement level, adhesion, and contractility. As a consequence, some physical characteristics of the tumor environment, such as the confinement imposed by the surrounding tissue on the growing tumor, might be enough to trigger a migratory switch toward a faster and more invasive mode of migration. Such an increase in the migration and invasion capacity of tumor cells has been observed using encapsulated spheroids as a model for tumor growth under confinement (Alessandri et al., 2013). This phenomenon could also explain how inhibition of matrix-degrading enzymes such as MMPs could directly promote a switch to amoeboid migration by increasing the confinement of the growing tumor and/or of single migrating cells (Wolf et al., 2003). In conclusion, our work demonstrates an unexpected plasticity of migration of transformed mesenchymal-like cells when they are not in their physiological physical environment, suggesting that tumor cells could spontaneously adopt a large variety of strategies to escape primary tumors and invade tissues, without any specific genetic alteration.

## EXPERIMENTAL PROCEDURES

In brief (see [Supplemental Information](#) for a detailed description of the methods), cell culture and experiments were performed in complete cell culture media adapted for each particular cell type. Drugs and RNAi treatments were performed following standard procedures. Cell confinement was performed using a home-made device (Le Berre et al., 2012) consisting of a suction cup made in polydimethylsiloxane (PDMS, RTV615, GE) used to press a confining coverslip bearing PDMS micro-spacers on top of the culture substrate plated with cells. Alternatively, a version of the cell confiner adapted to multi-well plates was used to perform multiple experiments in parallel. The molds for the PDMS micro-spacers were fabricated following standard photolithography procedures. The surface of the confining side was always treated with non-adhesive PLL-PEG (SuSoS), while the cell culture substrate was treated to obtain various degrees of cells adhesion using mixes of PLL-PEG and PLL-PEG-RGD. For automated cell tracking, cell nuclei were stained with Hoechst 33342. Image acquisition and image analysis were performed on workstations of the PICT-IBISA Lhomond Imaging facility of Institut Curie. All microscopes used for time-lapse recordings were equipped with an on-stage incubation chamber which maintained the temperature at 37°C and CO<sub>2</sub> concentration at 5% at all times. Image analysis was performed using Image J or MetaMorph software (Universal Imaging), or homemade routines under MATLAB (MathWorks). Analysis of cell trajectories was based on standard methods for analysis of persistent random walks. To measure single cell contractility index, cells were plated between two poly-acrylamide gels coated with beads and the cell dimensions as well as the gel deformation were measured by confocal microscopy. HeLa VL cells were always used as a reference for gel calibration. A minimal 1D theoretical model, based on the observations of the acto-myosin flows analysis, was adapted from Callan-Jones and Voituriez (2013) and Hawkins et al. (2011) to predict the mesenchymal to amoeboid transition.

## SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, six figures, and eight movies and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2015.01.007>.

## AUTHOR CONTRIBUTIONS

Y.-J.L. designed and carried out a majority of the experiments, performed data analysis, made the figures, and contributed to writing the paper. M.L.-B. carried out experiments, performed a majority of the data analysis, supervised the work, and wrote the paper. F.L. performed some TIRF experiments. P.M. carried out some data analysis. M.H. and T.T. performed immunoblotting. A.C.J. and R.V. made the physical model. M.P. designed experiments, supervised the work, and wrote the paper.

## ACKNOWLEDGMENTS

The authors acknowledge the Nikon Imaging Centre at the Institut Curie-CNRS and Vincent Fraisier from the PICT-IBISA Lhomond Imaging facility of Institut Curie for their help with the microscopy, and Timo Betz for kindly sharing his software for PIV analysis. We thank Stéphanie Miserey-Lenkei and Nicolas Carpi for their help with western blots. Vivatech and ARC are acknowledged for their financial support to Y.-J.L. and M.L.B., ANR-09-PIRI-0027 and InNaBioSanté to R.V. and M.P., and ERC 311205-PROMICO to M.P.

Received: May 2, 2014

Revised: October 31, 2014

Accepted: December 31, 2014

Published: February 12, 2015

## REFERENCES

- Alessandri, K., Sarangi, B.R., Gurchenkov, V.V., Sinha, B., Kießling, T.R., Fetler, L., Rico, F., Scheuring, S., Lamaze, C., Simon, A., et al. (2013). Cellular capsules as a tool for multicellular spheroid production and for investigating the mechanics of tumor progression in vitro. Proc. Natl. Acad. Sci. USA 110, 14843–14848.
- Barnhart, E.L., Lee, K.-C., Keren, K., Mogilner, A., and Theriot, J.A. (2011). An adhesion-dependent switch between mechanisms that determine motile cell shape. PLoS Biol. 9, e1001059.
- Bergert, M., Chandradoss, S.D., Desai, R.A., and Paluch, E. (2012). Cell mechanics control rapid transitions between blebs and lamellipodia during migration. Proc. Natl. Acad. Sci. USA 109, 14434–14439.
- Callan-Jones, A.C., and Voituriez, R. (2013). Active gel model of amoeboid cell motility. New J. Phys. 15, 025022.
- Carragher, N.O., Walker, S.M., Scott Carragher, L.A., Harris, F., Sawyer, T.K., Brunton, V.G., Ozanne, B.W., and Frame, M.C. (2006). Calpain 2 and Src dependence distinguishes mesenchymal and amoeboid modes of tumour cell invasion: a link to integrin function. Oncogene 25, 5726–5740.
- DiMilla, P.A., Stone, J.A., Quinn, J.A., Albelda, S.M., and Lauffenburger, D.A. (1993). Maximal migration of human smooth muscle cells on fibronectin and type IV collagen occurs at an intermediate attachment strength. J. Cell Biol. 122, 729–737.
- Feng, J., Ito, M., Ichikawa, K., Isaka, N., Nishikawa, M., Hartshorne, D.J., and Nakano, T. (1999). Inhibitory phosphorylation site for Rho-associated kinase on smooth muscle myosin phosphatase. J. Biol. Chem. 274, 37385–37390.
- Friedl, P., and Wolf, K. (2003). Tumour-cell invasion and migration: diversity and escape mechanisms. Nat. Rev. Cancer 3, 362–374.
- Giampieri, S., Manning, C., Hooper, S., Jones, L., Hill, C.S., and Sahai, E. (2009). Localized and reversible TGFbeta signalling switches breast cancer cells from cohesive to single cell motility. Nat. Cell Biol. 11, 1287–1296.
- Hawkins, R.J., Piel, M., Faure-Andre, G., Lennon-Dumenil, A.-M., Joanny, J.F., Prost, J., and Voituriez, R. (2009). Pushing off the walls: a mechanism of cell motility in confinement. Phys. Rev. Lett. 102, 058103.
- Hawkins, R.J., Poincloux, R., Bénichou, O., Piel, M., Chavrier, P., and Voituriez, R. (2011). Spontaneous contractility-mediated cortical flow generates cell migration in three-dimensional environments. Biophys. J. 101, 1041–1045.
- Heuzé, M.L., Vargas, P., Chabaud, M., Le Berre, M., Liu, Y.-J., Collin, O., Solanès, P., Voituriez, R., Piel, M., and Lennon-Duménil, A.-M. (2013). Migration of

dendritic cells: physical principles, molecular mechanisms, and functional implications. *Immunol. Rev.* 256, 240–254.

Hung, W.-C., Chen, S.-H., Paul, C.D., Stroka, K.M., Lo, Y.-C., Yang, J.T., and Konstantopoulos, K. (2013). Distinct signaling mechanisms regulate migration in unconfined versus confined spaces. *J. Cell Biol.* 202, 807–824.

Ishihara, H., Martin, B.L., Brautigan, D.L., Karaki, H., Ozaki, H., Kato, Y., Fusetani, N., Watabe, S., Hashimoto, K., Uemura, D., et al. (1989). Calyculin A and okadaic acid: inhibitors of protein phosphatase activity. *Biochem. Biophys. Res. Commun.* 159, 871–877.

Lämmermann, T., and Sixt, M. (2009). Mechanical modes of ‘amoeboid’ cell migration. *Curr. Opin. Cell Biol.* 21, 636–644.

Lämmermann, T., Bader, B.L., Monkley, S.J., Worbs, T., Wedlich-Söldner, R., Hirsch, K., Keller, M., Förster, R., Critchley, D.R., Fässler, R., and Sixt, M. (2008). Rapid leukocyte migration by integrin-independent flowing and squeezing. *Nature* 453, 51–55.

Le Berre, M., Aubertin, J., and Piel, M. (2012). Fine control of nuclear confinement identifies a threshold deformation leading to lamina rupture and induction of specific genes. *Integr. Biol. (Camb)* 4, 1406–1414.

Lorentzen, A., Bamber, J., Sadok, A., Elson-Schwab, I., and Marshall, C.J. (2011). An ezrin-rich, rigid uropod-like structure directs movement of amoeboid blebbing cells. *J. Cell Sci.* 124, 1256–1267.

Malawista, S.E., de Boisfleury Chevance, A., and Boxer, L.A. (2000). Random locomotion and chemotaxis of human blood polymorphonuclear leukocytes from a patient with leukocyte adhesion deficiency-1: normal displacement in close quarters via chimneying. *Cell Motil. Cytoskeleton* 46, 183–189.

Paluch, E.K., and Raz, E. (2013). The role and regulation of blebs in cell migration. *Curr. Opin. Cell Biol.* 25, 582–590.

Panková, K., Rösel, D., Novotný, M., and Brábek, J. (2010). The molecular mechanisms of transition between mesenchymal and amoeboid invasiveness in tumor cells. *Cell. Mol. Life Sci.* 67, 63–71.

Petrie, R.J., Gavara, N., Chadwick, R.S., and Yamada, K.M. (2012). Nonpolarized signaling reveals two distinct modes of 3D cell migration. *J. Cell Biol.* 197, 439–455.

Petrie, R.J., Koo, H., and Yamada, K.M. (2014). Generation of compartmentalized pressure by a nuclear piston governs cell motility in a 3D matrix. *Science* 345, 1062–1065.

Poincloux, R., Collin, O., Lizárraga, F., Romao, M., Debray, M., Piel, M., and Chavrier, P. (2011). Contractility of the cell rear drives invasion of breast tumor cells in 3D Matrigel. *Proc. Natl. Acad. Sci. USA* 108, 1943–1948.

Renkawitz, J., Schumann, K., Weber, M., Lämmermann, T., Pflicke, H., Piel, M., Polleux, J., Spatz, J.P., and Sixt, M. (2009). Adaptive force transmission in amoeboid cell migration. *Nat. Cell Biol.* 11, 1438–1443.

Ruprecht, V., Wieser, S., Callan-Jones, A., Smutny, M., Morita, H., Sako, K., Barone, V., Ritsch-Marte, M., Sixt, M., Voituriez, R., and Heisenberg, C.-P. (2015). Cortical contractility triggers a stochastic switch to fast amoeboid cell motility. *Cell* 160, this issue, 674–686.

Sahai, E., and Marshall, C.J. (2003). Differing modes of tumour cell invasion have distinct requirements for Rho/ROCK signalling and extracellular proteolysis. *Nat. Cell Biol.* 5, 711–719.

Sanz-Moreno, V., and Marshall, C.J. (2010). The plasticity of cytoskeletal dynamics underlying neoplastic cell migration. *Curr. Opin. Cell Biol.* 22, 690–696.

Taddei, M.L., Giannoni, E., Comito, G., and Chiarugi, P. (2013). Microenvironment and tumor cell plasticity: an easy way out. *Cancer Lett.* 341, 80–96.

Tozluoğlu, M., Tournier, A.L., Jenkins, R.P., Hooper, S., Bates, P.A., and Sahai, E. (2013). Matrix geometry determines optimal cancer cell migration strategy and modulates response to interventions. *Nat. Cell Biol.* 15, 751–762.

Trache, A., and Meininger, G.A. (2008). Total internal reflection fluorescence (TIRF) microscopy. In *Current Protocols in Microbiology*, R. Coico, T. Kowalik, J. Quarles, B. Stevenson, and R. Taylor, eds. (John Wiley & Sons).

Wilson, C.A., Tsuchida, M.A., Allen, G.M., Barnhart, E.L., Applegate, K.T., Yam, P.T., Ji, L., Keren, K., Danuser, G., and Theriot, J.A. (2010). Myosin II contributes to cell-scale actin network treadmilling through network disassembly. *Nature* 465, 373–377.

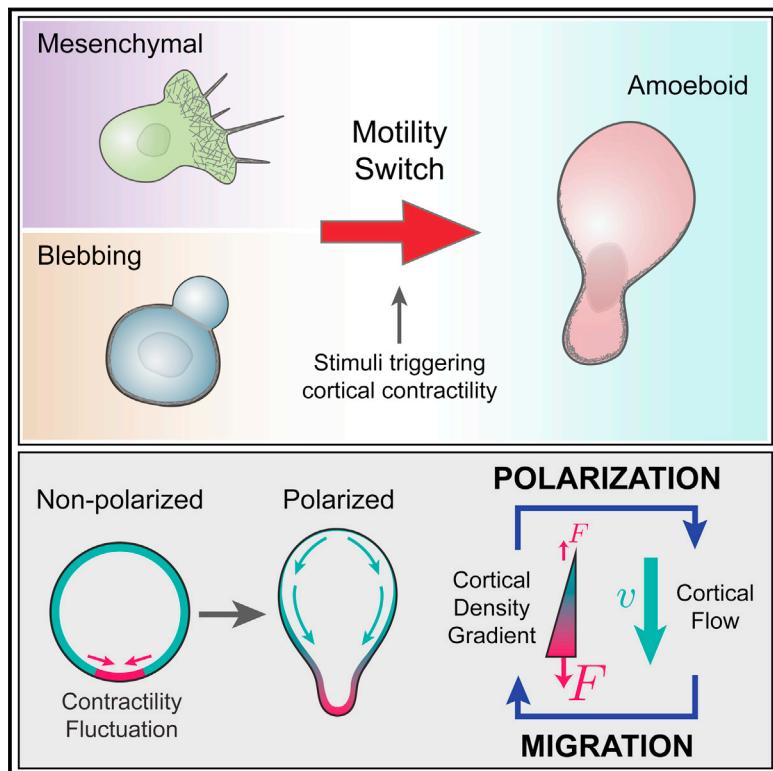
Wilson, K., Lewalle, A., Fritzche, M., Thorogate, R., Duke, T., and Charras, G. (2013). Mechanisms of leading edge protrusion in interstitial migration. *Nat. Commun.* 4, 2896.

Wolf, K., Mazo, I., Leung, H., Engelke, K., von Andrian, U.H., Deryugina, E.I., Strongin, A.Y., Bröcker, E.-B., and Friedl, P. (2003). Compensation mechanism in tumor cell migration: mesenchymal-amoeboid transition after blocking of pericellular proteolysis. *J. Cell Biol.* 160, 267–277.

Yoshida, K., and Inouye, K. (2001). Myosin II-dependent cylindrical protrusions induced by quinine in Dictyostelium: antagonizing effects of actin polymerization at the leading edge. *J. Cell Sci.* 114, 2155–2165.

# Cortical Contractility Triggers a Stochastic Switch to Fast Amoeboid Cell Motility

## Graphical Abstract



## Authors

Verena Ruprecht, Stefan Wieser, ...,  
Raphaël Voituriez,  
Carl-Philipp Heisenberg

## Correspondence

[ruprecht.verena@gmail.com](mailto:ruprecht.verena@gmail.com) (V.R.),  
[heisenberg@ist.ac.at](mailto:heisenberg@ist.ac.at) (C.-P.H.)

## In Brief

Cell-intrinsic fluctuation in cortical contraction forces triggers the switch in motility behavior that allows embryonic progenitor cells to acquire a fast and persistent migratory mode in confined 3D environments

## Highlights

- Embryonic progenitor cells transform into a prototypic amoeboid migration mode
- Contractility driven cortical network instabilities drive rapid cell polarization
- Cell polarization is maintained by a positive cortical feedback loop
- Cortical flows drive fast and persistent cell motility in confined 3D environments

# Cortical Contractility Triggers a Stochastic Switch to Fast Amoeboid Cell Motility

Verena Ruprecht,<sup>1,5,\*</sup> Stefan Wieser,<sup>1,3,5</sup> Andrew Callan-Jones,<sup>4</sup> Michael Smutny,<sup>1</sup> Hitoshi Morita,<sup>1</sup> Keisuke Sako,<sup>1</sup> Vanessa Barone,<sup>1</sup> Monika Ritsch-Marte,<sup>3</sup> Michael Sixt,<sup>1</sup> Raphaël Voituriez,<sup>2</sup> and Carl-Philipp Heisenberg<sup>1,\*</sup>

<sup>1</sup>Institute of Science and Technology Austria, 3400 Klosterneuburg, Austria

<sup>2</sup>Laboratoire Jean Perrin and Laboratoire de Physique Théorique de la Matière Condensée, CNRS/Université Pierre et Marie Curie, 75005 Paris, France

<sup>3</sup>Division of Biomedical Physics, Innsbruck Medical University, 6020 Innsbruck, Austria

<sup>4</sup>Laboratoire Matière et Systèmes Complexes, CNRS/Université Paris-Diderot, UMR 7057, 75204 Paris Cedex 13, France

<sup>5</sup>Co-first author

\*Correspondence: ruprecht.verena@gmail.com (V.R.), heisenberg@ist.ac.at (C.-P.H.)

<http://dx.doi.org/10.1016/j.cell.2015.01.008>

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

## SUMMARY

**3D amoeboid cell migration is central to many developmental and disease-related processes such as cancer metastasis. Here, we identify a unique prototypic amoeboid cell migration mode in early zebrafish embryos, termed stable-bleb migration.** Stable-bleb cells display an invariant polarized balloon-like shape with exceptional migration speed and persistence. Progenitor cells can be reversibly transformed into stable-bleb cells irrespective of their primary fate and motile characteristics by increasing myosin II activity through biochemical or mechanical stimuli. Using a combination of theory and experiments, we show that, in stable-bleb cells, cortical contractility fluctuations trigger a stochastic switch into amoeboid motility, and a positive feedback between cortical flows and gradients in contractility maintains stable-bleb cell polarization. We further show that rearward cortical flows drive stable-bleb cell migration in various adhesive and non-adhesive environments, unraveling a highly versatile amoeboid migration phenotype.

## INTRODUCTION

Migrating cells show a versatile repertoire of migration modes with remarkable plasticity, allowing them to switch between different migration strategies in response to changing environmental conditions and activation of distinct molecular pathways (Friedl and Alexander, 2011). In order to migrate, cells need to establish an axis of polarity prior to movement. This polarity ultimately manifests itself in a polarized architecture of the actomyosin network, which in turn drives cell locomotion through different mechanical principles: in mesenchymal migration, the cortical actomyosin network facilitates unidirectional movement via polarized actin polymerization at the leading edge, combined with myosin-based contraction at the cell rear to disassemble adhesion sites. Amoeboid cells, in contrast, show hetero-

geneous shape and motility characteristics with actin-based protrusions, such as lamellipodia and pseudopodia and contraction-mediated protrusions, such as cellular blebs (Lämmermann and Sixt, 2009). Recent studies have suggested that propulsive forces in amoeboid cells are generated by cortical contractility and retrograde cortical flows (Blaser et al., 2006; Poincloux et al., 2011; Shih and Yamada, 2010), allowing movement even in the absence of specific adhesive coupling to the environment (Lämmermann and Sixt, 2009; Tozluoglu et al., 2013).

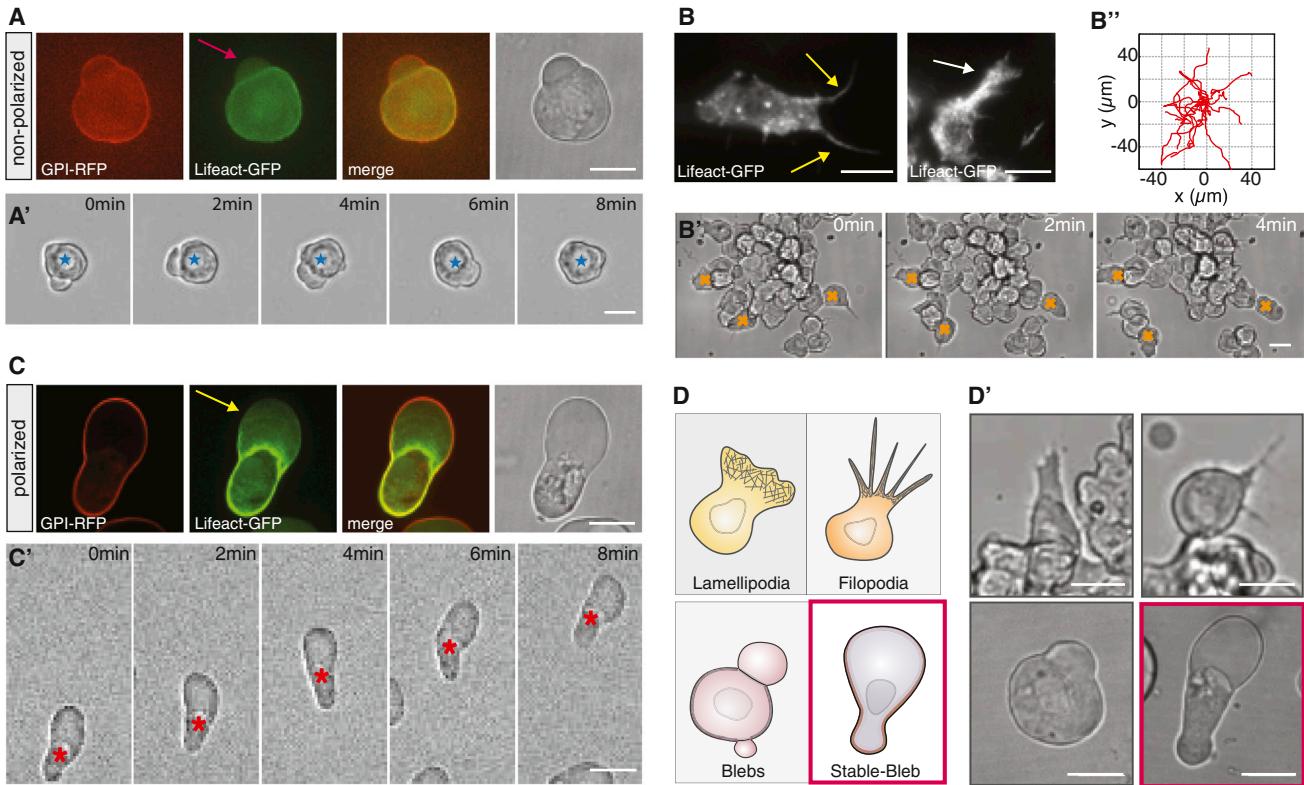
During zebrafish gastrulation, progenitor cells become motile and undergo extensive migration to form the ectoderm, mesoderm, and endoderm germ layers. While ectodermal progenitors assemble in a pseudo-epithelial cell layer, mesodermal and endodermal (mesendodermal) progenitor cells display a highly motile mesenchymal phenotype with a mixture of lamellipodia and bleb-like protrusions (Row et al., 2011). Interfering with the ratio of those protrusion types has been shown to lower the directionality but not the speed of their migration (Díz-Muñoz et al., 2010). Besides mesendodermal progenitors, primordial germ cells (PGCs) also undergo extensive migration during gastrulation but nearly exclusively use bleb-like protrusions for their migration (Blaser et al., 2006). Although using different protrusion types, migration speed and directionality of PGCs and mesendodermal progenitors appear surprisingly similar (Blaser et al., 2006; Díz-Muñoz et al., 2010), raising questions as to the choice and benefit of certain protrusion types over others for the migration of the different progenitor cell types during gastrulation.

Here, we have studied different migration phenotypes during zebrafish gastrulation and identified a cortical contractility-mediated cell-intrinsic motility switch to fast amoeboid migration in 3D environments, which we termed stable-bleb migration.

## RESULTS

### Identification of Basic Migration Modes in Zebrafish Germ Layer Progenitor Cells

To study the emergence of migration competence in early germ layer progenitor cells, we aimed at developing *in vitro* assays to investigate the complex range of migration behaviors observed *in vivo* under controlled conditions with a minimal set of defined



**Figure 1. Zebrafish Germ Layer Progenitor Cells Exhibit Distinct Cell Migration Modes In Vitro**

(A) Bright-field (BF) and fluorescence images of blebbing progenitor cells cultured in serum-free medium in confinement with GPI-RFP (membrane, red) and Lifeact-GFP (cortex, green). Red arrow marks cellular bleb. Scale bar represents 20  $\mu$ m.

(A') BF time-lapse images of non-motile progenitors related to culture conditions in (A). Blue asterisk indicates cell center. Scale bar represents 20  $\mu$ m.

(B) TIRFM images of Lifeact-GFP in mesendodermal progenitors cultured on 2D substrates. Arrows mark filopodia (yellow) and lamellipodia (white). Scale bar represents 10  $\mu$ m.

(B') BF time-lapse images related to culture conditions in (B). Orange asterisks indicate exemplary migrating cells. Scale bar represents 20  $\mu$ m.

(B'') Representative tracks of motile mesendodermal cells over 20 min (time lag 30 s, n = 15 cells).

(C) BF and fluorescence images of polarized progenitor cells with GPI-RFP (red) and Lifeact-GFP (green) cultured in 20% serum in confinement. Yellow arrow points at the bleb-like protrusion front. Scale bar represents 20  $\mu$ m.

(C') BF time-lapse images of motile stable-bleb cells cultured as described in (C). Red asterisk indicates cell movement. Scale bar represents 20  $\mu$ m.

(D and D') Sketch and BF images summarizing different migration phenotypes of embryonic progenitor cells in vitro. All progenitor cells were obtained from embryos at sphere stage and cultured on Fibronectin-coated substrates. Scale bar represents 20  $\mu$ m.

See also Movie S1 and Extended Experimental Procedures.

environmental parameters. Early progenitor cells placed on 2D substrates displayed a characteristic blebbing morphology that can also be observed in early blastula stage embryos *in vivo* (Diz-Muñoz et al., 2010). Notably, those blebbing cells failed to migrate irrespective of adhesive substrate coating with extracellular matrix (ECM) components, such as Laminin or Fibronectin (Figure 1A; Movie S1 available online). However, when progenitor cells were induced to be of mesodermal or mesendodermal origin and placed on Fibronectin-coated substrates, they formed a characteristic mixture of lamellipodia and filopodia (Figure 1B) and underwent collective migration with similar speed ( $<\!v_{Cell}>$  =  $3.8 \pm 0.3 \mu\text{m}/\text{min}$ ) to their movement *in vivo* (Figure 1B'; Movie S1).

Strikingly, when adding serum to the culture medium, we observed unexpected changes in progenitor cell architecture with cells displaying a highly polarized cell morphology charac-

terized by a stable pear-like shape and a large spherical protrusion front (Figure 1C). These cells, which we termed stable-bleb cells, were non-motile when plated on adhesive 2D substrates, but became highly migratory in confined environments (Figure 1C'; Movie S1). Thus, in the presence of serum, progenitors transformed into a novel migration mode *in vitro*, clearly distinct from previously described lamellipodia-, filopodia-, and bleb-based migration types (Figure 1D).

### Changes in Cortical Architecture Precede Stable-Bleb Cell Polarization

To investigate the molecular basis underlying this transformation, we first asked which serum components are involved. While adding serum growth factors, such as fibroblast growth factor (FGF), platelet-derived growth factor (PDGF), or epidermal growth factor (EGF), to the culture medium was not sufficient

to induce stable-bleb cells (Figure S1A), we observed a fast and robust transformation when lysophosphatidic acid (LPA) was added (Figures 2A and S1B), a serum phospholipid known to activate cortical contractility via the Rho/Rock pathway (Mills and Moolenaar, 2003).

Polarization was stable in the presence of LPA, but cells switched rapidly back into their original blebbing behavior upon dilution of LPA from the culture medium, indicating that LPA-mediated transformation of progenitors into stable-bleb cells is a reversible process (Figure S1C). Furthermore, cell polarization was observed in suspension at homogenous levels of LPA (Movie S2) suggesting that cell transformation occurs in the absence of extrinsic gradients of biochemical or environmental polarity cues and arguing in favor of a cell intrinsic stochastic polarization mechanism induced by LPA.

Surprisingly, transformation into stable-bleb cells could be triggered in progenitor cells from blastula and gastrula stages irrespective of their primary cell fate (Figure 2B) as determined by dissociating cells from *Tg(mezzo:eGFP)* transgenic embryos expressing eGFP in mesendodermal progenitor cells. We thus reasoned that the motile switch occurred independently of a cell-type-specific genetic module, but might rather be associated with a generic polarization mechanism retrievable in these cells.

Given that LPA is known to activate cortical contractility via the Rho/Rock pathway, we hypothesized that a mechanical polarization mechanism of the cell cortex could trigger the transformation of progenitor cells into stable-bleb cells (Carvalho et al., 2013). To test this hypothesis, we treated stable-bleb cells with the Rho kinase inhibitor Y-27632 or the myosin-II inhibitor Blebbistatin (Figure 2C). Treated cells lost their characteristic polarization, supporting a critical role for Rho/Rock-mediated cortical contractility in driving stable-bleb cell transformation.

To address the effect of LPA stimulation on stable-bleb cell polarization, we monitored the distribution of cortical markers in the presence of different LPA concentrations. Myosin II rapidly redistributed to the cell cortex upon LPA stimulation ( $t_{1/2} \sim 20$  s, Figure 2D), and rising LPA concentrations were associated with increasing cortical accumulations of myosin II, eventually saturating at a threshold level at which progenitors transformed into stable-bleb cells (Figures 2E and 2F). Importantly, the accumulation of myosin II at the cortex was up to seven times larger than the accumulation of actin, indicating that myosin II-mediated cortical contractility was specifically elevated upon LPA stimulation (Figures 2G and S1D). Furthermore, bleb sizes in dissociated progenitor cells closely scaled with the concentrations of LPA in the medium (Figures 2E and 2H), suggesting that LPA upregulates cortical contractility and enhances bleb expansion by increasing intracellular pressure.

Interestingly, we observed that strong spatial confinement of progenitor cells in serum-free medium also led to elevated levels of myosin II accumulation at the cortex, larger blebs, and stochastic transformation into stable-bleb cells (Figures 2I–2K). In mesenchymal cells, this transformation was typically preceded by an increase in the formation of cellular blebs at the expense of lamellipodia and filopodia (Figure 2L). Together, these results imply that increasing levels of myosin-II mediated cortical contractility by either biochemical or physical stimuli can trigger

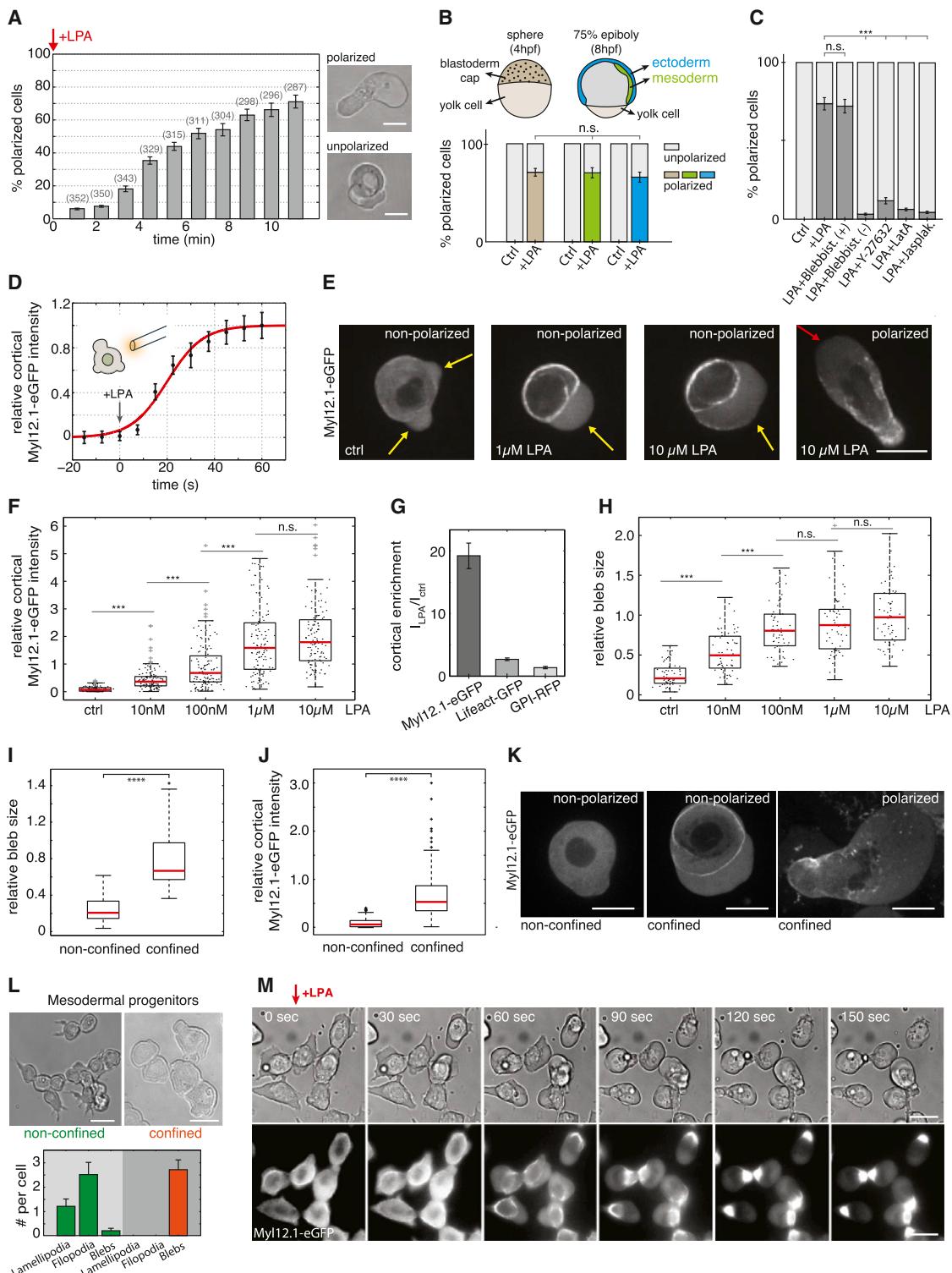
a rapid and reversible amoeboid transformation of non-polarized and mesenchymal progenitors (Figure 2M; Movie S3) that is accompanied with extensive remodeling of cortical architecture and myosin II localization in those cells. Notably, a similar mesenchymal-to-amoeboid transition (MAT) has been observed in various culture cell types upon compression (see Liu et al. [2015] in this issue of Cell), suggesting that contractility-induced MAT is a universal mechanism retrievable in various primary and culture cell types.

### A Contraction-Based Polarization Mechanism Drives Stable-Bleb Cell Transformation

Based on these experimental results, we hypothesized that rapid changes in cortical architecture by increasing levels of cortical contractility mediate a cell intrinsic transformation into stable-bleb cells. To address this hypothesis, we developed a theoretical description of the polarization process by modeling cortex structures and associated cell shapes as a function of cortical contractility (see Extended Experimental Procedures). In this model, the cell cortex was treated as a thin spherical shell of compressible active gel (Callan-Jones and Voituriez, 2013; Hawkins et al., 2011) (Figure S2A). Assuming homogenous bulk polymerization and depolymerization rates, we precluded a predetermined symmetry break of the cortical layer. Instead, we treated blebbing events in unpolarized progenitor cells as random spatial fluctuations in cortical contractility (Figure 3A), an assumption supported by measurements of the spatial extent (bleb size) and amplitude of cortical instabilities (myosin II enrichment) during cell blebbing at increasing LPA levels (Figures S2B and S2C). Notably, our model predicted that above a threshold level of cortical contractility these local instabilities would not relax but be amplified, eventually transforming an initially homogenous cortex layer into a polarized steady state characterized by a stable cortical density gradient along with retrograde cortical flows (Figures 3B and S2D).

Indeed, time-lapse imaging of myosin II localization during cell polarization in LPA stimulated progenitor cells confirmed a rapid myosin II re-localization to the cortex and strongly increasing bleb sizes prior to polarization, consistent with elevated levels of cortical contractility fluctuations and an increasing imbalance in contractile forces between the bleb and cellular cortical network (Figure 3C; Movie S4). Stabilization of blebs was initiated by the appearance of directional cortical flows from the weak actomyosin-network assembling underneath the bleb toward the cell cortex, which resulted in a failure of bleb retraction and ultimately led to stable cell polarization (Movie S4). To test whether ectopically increased levels of contractility can trigger cell polarization, we generated a localized LPA diffusion gradient by a micropipette (Figure 3D). Progenitor cells in the presence of this LPA gradient rapidly transformed into stable-bleb cells with their contractile back oriented toward the source of LPA (Figures 3E, 3F, and S2E). This suggests that external gradients of LPA can trigger directional cell polarization by inducing an asymmetric contraction of the cortical cytoskeleton in agreement with theoretical modeling.

We next asked how stable-bleb cells maintain their polarity. According to our theoretical model, we expected that cell polarization in stable-bleb cells could potentially be maintained by a

**Figure 2. LPA Induces Cell Polarization to Stable-Bleb Cells In Vitro**(A) Percentage of polarized stable-bleb cells over time upon stimulation with 100  $\mu$ M LPA.(B) Percentage of polarized stable-bleb cells from blastula stage embryos (brown,  $n = 287$ ) or 75% epiboly stage *Tg(mezeo:eGFP)* transgenic embryos with mesendodermal (green,  $n = 183$ ) and ectodermal progenitors (blue,  $n = 152$ ) after 10 min in the presence of 100  $\mu$ M LPA (+LPA) or DMEM-F12 culture medium alone (Control [Ctrl] at blastula stage [ $n = 321$ ], mesendodermal [ $n = 286$ ], and ectodermal [ $n = 251$ ]).

(legend continued on next page)

positive feedback loop between cortical contractility gradients and the presence of cortical flows (Bois et al., 2011; Hawkins et al., 2011). In this positive feedback, cortical flows reinforce density gradients, in particular rearward localization of myosin II, and thereby reinforce contractility gradients that drive cortical flow toward the contractile region (Figure 3G). High resolution TIRF imaging of cortical actin and myosin II in polarized progenitor cells confirmed the presence of stable cortical density gradients toward the cell rear and revealed a low density actomyosin network in the spherical protrusion front (Figures 4A, 4B, S3A, and S3B). This sparse actomyosin meshwork was reminiscent of a bleb-like membrane blister but, unlike blebs, was accompanied by an unusually fast and continuous cortical actomyosin flow, referred to as cortical flow in the following, with maximal flow speeds up to 150  $\mu\text{m}/\text{min}$  in the very cell front (Figures 4A, 4C, S3C, and S3D). Measurement of cortical flows along with cortical density profiles allowed for calculating cortex flux and cortex turnover rate (Figure 4D), indicating net polymerization in the spherical protrusion front and de-polymerization toward the rear. As a continuous rearward cortex flux requires permanent cortex turnover, stable-bleb cell polarization was rapidly lost upon treatment with the G-actin sequestering drug Latrunculin A or Jasplakinolide, a drug that interferes with cortex turnover (Figures 2C and S1E; Movie S2). Moreover, treatment with the myosin II inhibitor Blebbistatin also reversed cell polarization (Figure 2C), indicating that cortical flow in combination with a gradient in contractility is critical for maintaining stable-bleb cell polarization over time. In contrast, inhibition of CDC42 (ML-141) or PI3Kinase (L-294002), previously shown to be required for mesendodermal progenitor cell migration *in vivo* (Montero et al., 2003), did not affect stable-bleb cell polarization (Figure S2F), supporting the concept that stable-bleb cell motility is unrelated to actin driven protrusion types such as lamellipodia or filopodia. Collectively, our results support a simple mechanical model of stochastic cell polarization based on the amplification of local fluctuations in cortical contractility and a positive feedback mechanism between contractility gradients and continuous cortical flows maintaining polarity in stable-bleb cells (Figure 3H).

To determine how the polarized cortical scaffold in stable-bleb cells connects to their characteristic shape, we expanded our theoretical model to calculate cell shapes based on deforma-

tions of the cytoplasm due to compressive and shear stresses at the cytoplasmic surface arising from cortical tension and flow-friction (Behrndt et al., 2012) (Figures S3E and S3F). The cytoplasm was modeled as a viscoelastic medium with residual elastic response at long times (Tinevez et al., 2009), which is the regime of interest for discussing steady-state shapes. While measured actomyosin density gradients serve as a readout of cortical contractility, cortical tension in general is related not only to contractility but also to viscous forces arising from flow (Mayer et al., 2010) and flow-induced filament alignment (Salbreux et al., 2009). Using experimentally determined cortical density profiles we calculated shapes for two different cases: In the first case, we assumed an isotropic tension resulting in flattened, bulbous cell shapes not matching our experimental observations (Figure 4G, bottom). In the second case, we included viscous forces leading to tension reduction along the flow direction. Such tension anisotropy resulted in elongated pear shapes closely mirroring the observed shapes of stable-bleb cells (Figure 4G, top). In support of our theoretical predictions, we observed a preferential alignment of actin filaments perpendicular to the front-to-rear axis of stable-bleb cells, highlighting the presence of anisotropic tension within the cortex (Figures S3I–S3K). Finally, we tested whether stable-bleb cell shapes depend on cortical viscosity and cell stiffness (Figures 4H and S3G) and used osmotic shock experiments to alter intracellular pressure and, consequently, cell stiffness. Consistent with our theoretical predictions, we found that decreasing intracellular pressure resulted in more elongated cell shapes (Figure S3H). Together, our theoretical considerations suggest that the characteristic pear-shape morphology of stable-bleb cells is generated by anisotropic tension in the cell cortex, which acts as a belt-like constriction unit applying compressive forces normal to the axis of polarization.

### Retrograde Cortical Flows Drive Stable-Bleb Cell Migration

To study the migratory behavior of polarized stable-bleb cells, we monitored their motility in different environments (Figure 5A). Stable-bleb cells only weakly adhered to and failed to migrate on 2D adhesive substrates coated with ECM components, such as Laminin and Fibronectin. Fluorescence imaging of the

(C) Percentage of polarized stable-bleb cells cultured in DMEM-F12 medium (Ctrl, n = 272), 100  $\mu\text{M}$  LPA (+LPA, n = 219) or 30  $\mu\text{M}$  LPA supplemented with 10  $\mu\text{M}$  Blebbistatin (+) (n = 167), 10  $\mu\text{M}$  Blebbistatin (−) (n = 143), 10  $\mu\text{M}$  Y-27632 (n = 149), 100 nM Latrunculin A (LatA, n = 134), and 500 nM Jasplakinolide (Jasplak) (n = 83).

(D) Recruitment of myosin II to the cortex of progenitor cells over time after application of 10  $\mu\text{M}$  LPA by a micropipette. Data (black, mean  $\pm$  SEM) and sigmoid fit function (red) with half time  $t_{1/2} = 20$  s (n = 30).

(E) Myosin II localization in progenitor cells cultured in DMEM-F12 medium alone (Ctrl), 1  $\mu\text{M}$  or 10  $\mu\text{M}$  LPA. Arrows mark blebs (yellow) and the cortex-depleted protrusion front of stable-bleb cells (red).

(F) Boxplot of relative cortical myosin II fluorescence intensities for increasing LPA levels in unpolarized progenitor cells (each n = 120).

(G) Cortical enrichment of Myl12.1-eGFP (n = 60), Lifeact-GFP (n = 60), and GPI-RFP (n = 33) in non-polarized progenitor cells treated with 10  $\mu\text{M}$  LPA. Cortical fluorescence intensity values  $I_{\text{LPA}}$  are normalized to reference values without LPA stimulation  $I_{\text{Ctrl}}$ .

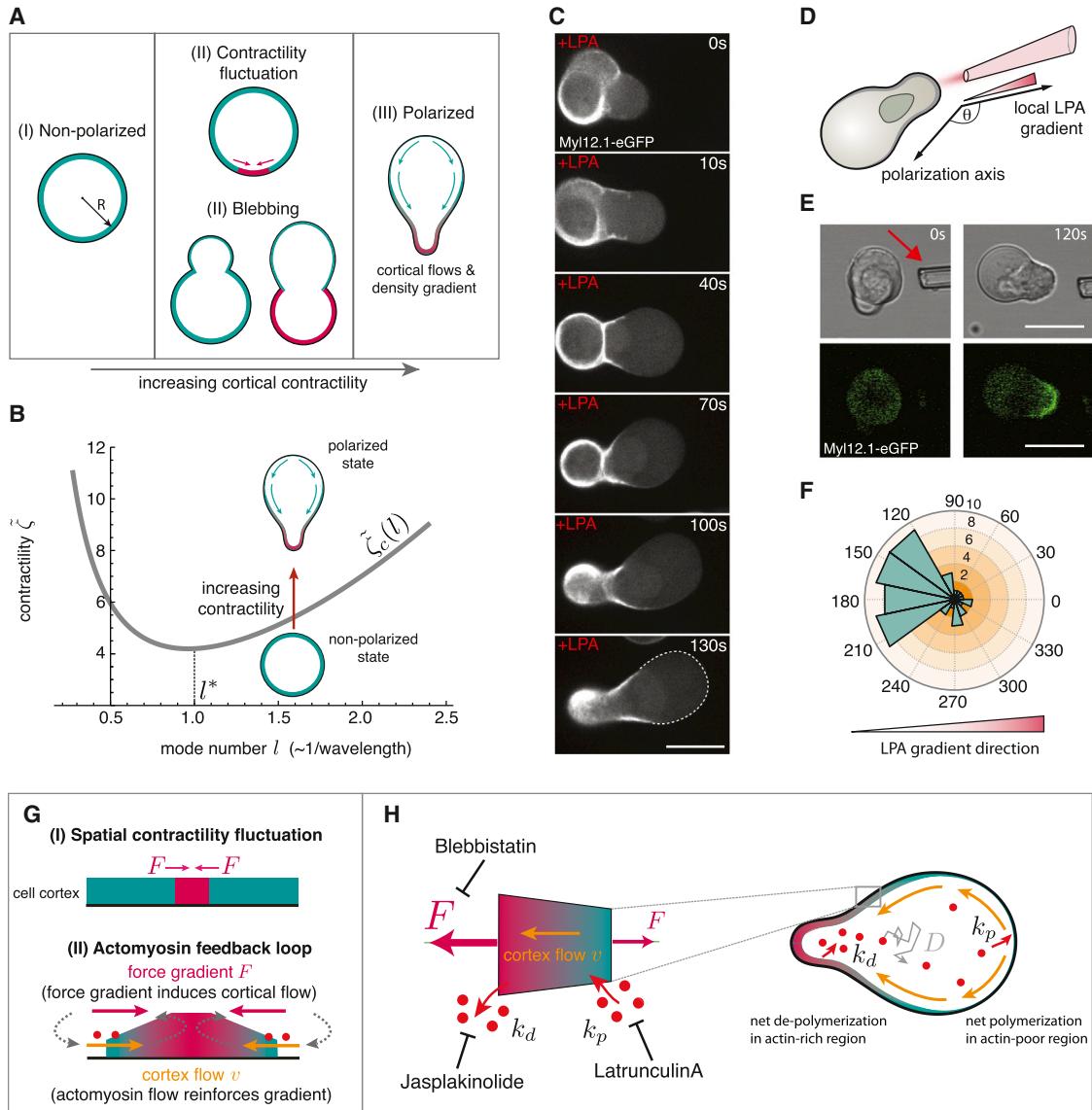
(H) Relative bleb sizes for increasing LPA levels (each n = 120).

(I and J) Relative bleb sizes (I) and relative cortical myosin II fluorescence intensities (J) under non-confined and confined conditions.

(K) Exemplary fluorescence images of non-polarized and polarized cells expressing Myl12.1 in non-confined and confined conditions.

(L) BF images of mesodermal progenitors cultured in non-confined and confined environments (top) and quantification of cellular protrusions (bottom).

(M) Time-lapse BF and fluorescence images of Myl12.1-eGFP localization in mesodermal progenitor cells cultured on a 2D substrate coated with Fibronectin upon LPA stimulation. n.s., not significant, \*\*\*p < 0.001. Boxplots show single cell data points (black) and median (red). All cells were obtained from embryos at sphere stage, despite (B), and cultured in suspension on a passivated substrate despite (L and M). n, number of cells. All scale bars represent 20  $\mu\text{m}$ . See also Figure S1, Movies S2 and S3, and Extended Experimental Procedures.



**Figure 3. Cortical Contractility Fluctuations Drive Polarization to Stable-Bleb Cells**

(A) Sketch of polarization from a non-polarized state (left) to a polarized stable-bleb state (right). The cell cortex is shown in green and highly contractile cortex areas in red. Red arrows denote a cortical force imbalance due to local fluctuations in cortical contractility. Blebs induce spatial inhomogeneities in cortical contractility of increasing strength for larger bleb sizes. Green arrows indicate cortical flows.

(B) Stability diagram illustrating the switch from a non-polarized to a polarized state. The curve represents the values of cortical contractility  $\zeta$ , depending on the density fluctuation mode number  $l$ . The mode number,  $l$ , is inversely proportional to the wavelength (in units of initial cell radius) of the fluctuation.

(C) Myosin II localization in a transforming progenitor cell cultured in suspension with 30  $\mu$ M LPA. Scale bar represents 20  $\mu$ m.

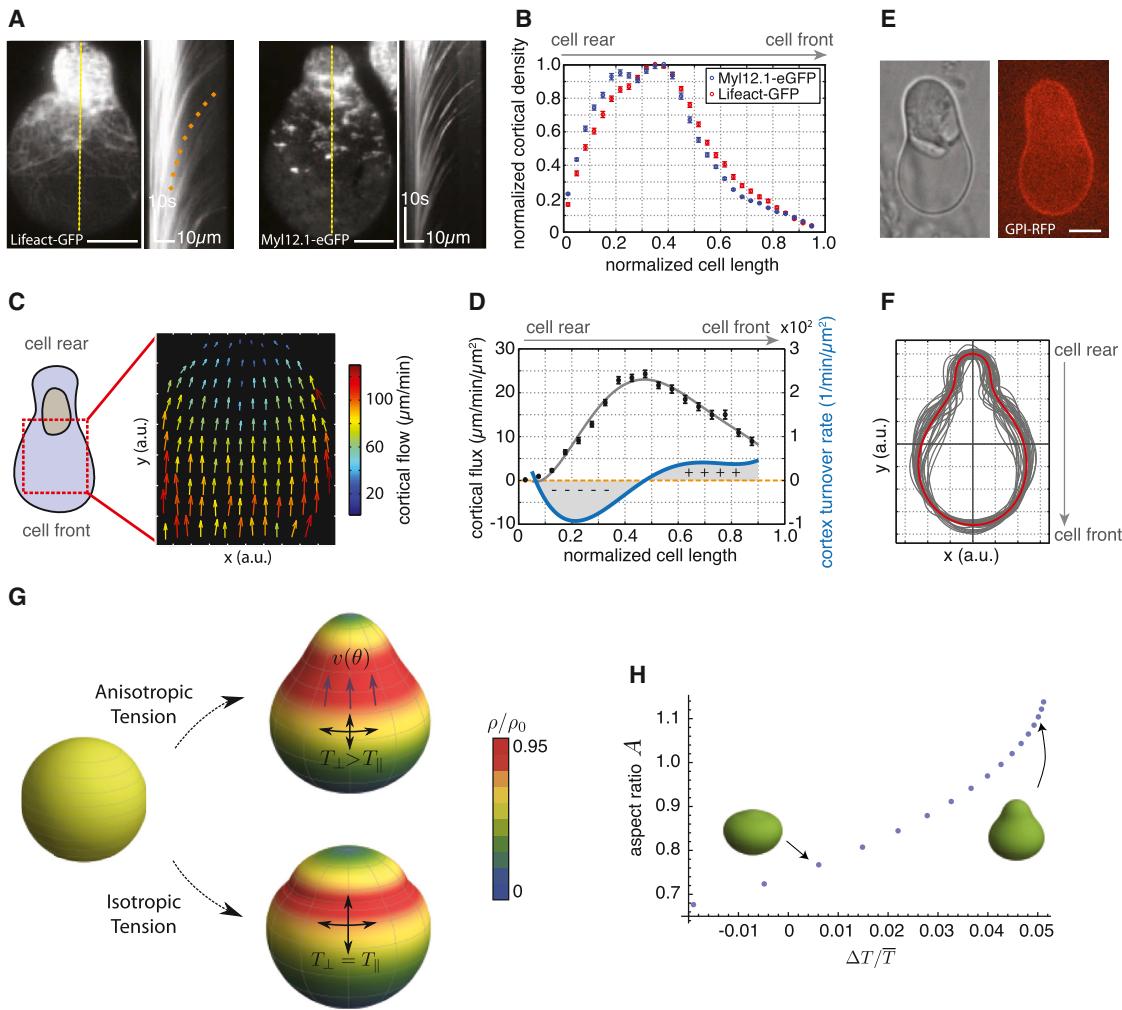
(D) Sketch of the local LPA application experiment.

(E) BF time-lapse images and Myl12.1-eGFP localization during progenitor cell polarization in a local 100  $\mu$ M LPA diffusion gradient set by a micropipette (red arrow).

(F) Histogram of measured cell polarization orientations ( $n = 32$ ).

(G and H) Schematic illustration of the positive feedback loop between cortical actomyosin density gradients and rearward cortical flows for the maintenance of stable-bleb cell polarity. Local fluctuations in cortical contractility induce a flow toward the contractile region along with mass transport of actin and myosin (red dots) that reinforces the initial instability (G). To maintain a continuous cortical flow in polarized cells actin turnover (with rate constants  $k_p$  and  $k_d$ ) and diffusion of free actin and myosin (with diffusion constant  $D$ ) is required (H). Polarization is lost upon pharmacological inhibition of myosin II contractility (by Blebbistatin), actin polymerization (by Latrunculin-A) and actin de-polymerization (by Jasplakinolide).

See also Figure S2, Movies S2 and S4, and Extended Experimental Procedures.



**Figure 4. Cortical Architecture Determines Cell Shape of Stable-Bleb Cells In Vitro**

(A) TIRFM image showing Lifeact-GFP (left) and myosin II localization (right) in isolated stable-bleb cells with corresponding kymograph data along yellow lines. Orange dotted line indicates the cortical flow profile.

(B) Average actin and myosin II density profiles obtained from culture conditions in (A) ( $n = 30$ ).

(C) Average 2D cortical flow map in the spherical protrusion front of stable-bleb cells ( $n = 3$ ). The red dashed square highlights the analyzed cell area.

(D) Cortical flux data (black, mean  $\pm$  SEM) calculated from average cortical density (B) and average cortical flow profiles (Figure S3C) with polynomial fit (gray) and cortex turnover rate (blue) calculated as the derivative of the cortical flux ( $n = 30$ ).

(E) BF and fluorescence cross-sectional image of an isolated stable-bleb cell expressing GPI-RFP.

(F) Average (red) and single cell shapes (gray) of polarized stable-bleb cells ( $n = 30$ ).

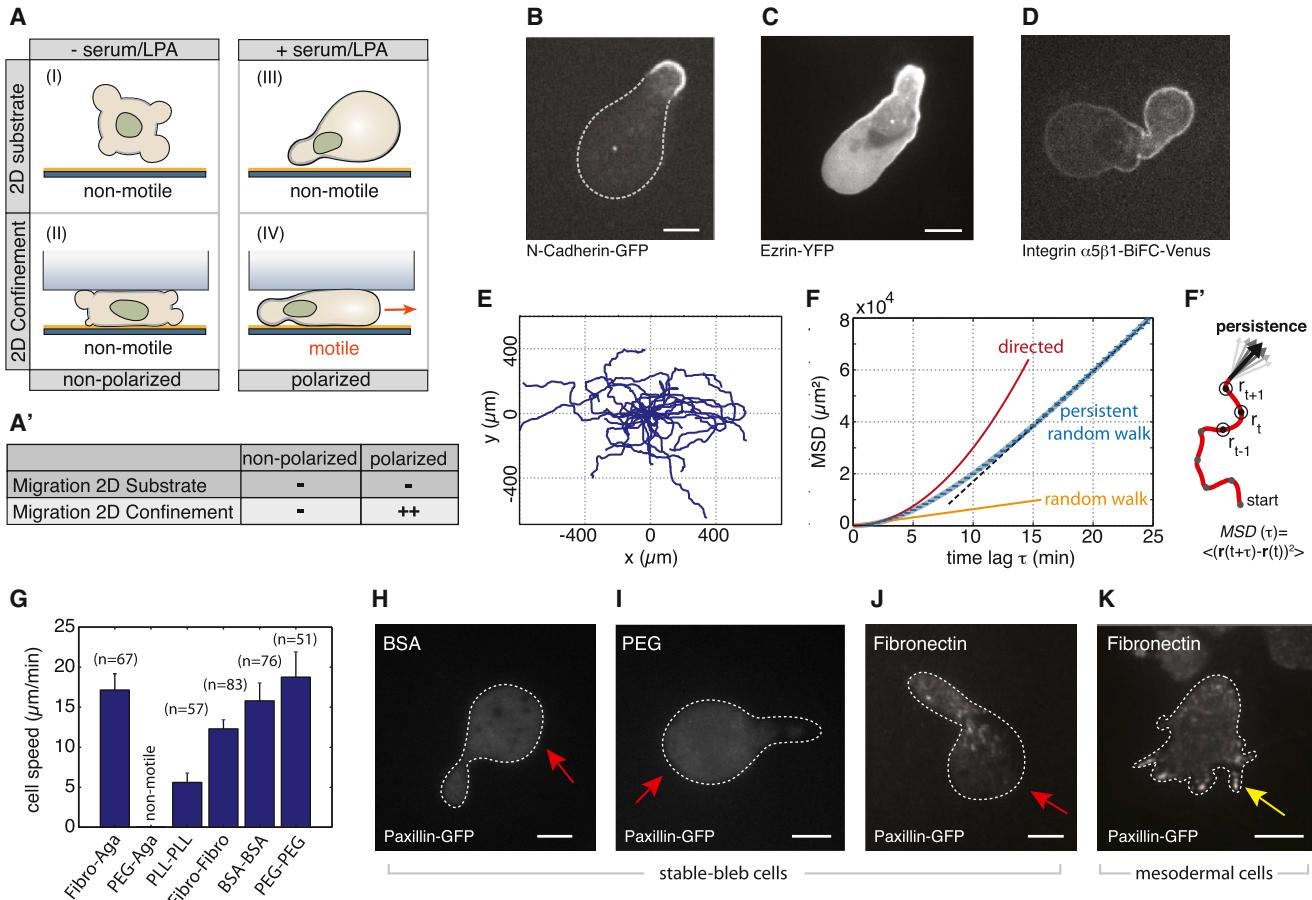
(G) Theoretical predictions on polarized cell shapes generated by cortical tension components parallel  $T_{\parallel}$  and perpendicular  $T_{\perp}$  to the polarization axis. Top: anisotropic tension ( $T_{\parallel} < T_{\perp}$ ). Bottom: isotropic tension ( $T_{\parallel} = T_{\perp}$ ). Color bar represents normalized cortical density distributions. Magenta arrows indicate the direction of cortical flow  $v(\theta)$  in polar direction. Parameter values are provided in the [Extended Experimental Procedures](#).

(H) Phase diagram of cell shapes with cell aspect ratio  $A$  as a function of tension anisotropy. Elongated, pear shapes are obtained for higher anisotropy. All cells were cultured on a PEG-coated substrate either with (A–D) or without confinement (E and F). Scale bars represent 10  $\mu$ m.

See also Figure S3 and [Extended Experimental Procedures](#).

subcellular distribution of molecules associated with coupling the cortex to the surrounding substrate, such as N-Cadherin, Ezrin, and Integrin revealed a highly asymmetric accumulation toward the cell rear, while the bleb-like protrusion front was largely devoid of those molecules (Figures 5B and 5C). This indicates only weak adhesive forces in the front that are insufficient for forward locomotion on 2D substrates. To increase frictional coupling of migrating cells, we thus created a planar 2D confine-

ment by placing cells between a coated glass substrate and an agarose layer. We found that stable-bleb cells plated on Fibronectin-coated substrates under confinement displayed exceptionally fast and directional movements with average cell speed  $\langle v_{cell} \rangle = 17.1 \pm 0.6 \mu\text{m}/\text{min}$  ( $n = 67$  cells) and persistence length of approximately four cell diameters (Figures 5E and 5F), best described by a persistent random walk model over long time-scales (Figure 5F).



**Figure 5. Migration Characteristics of Stable-Bleb Cells In Vitro**

(A) Illustration of progenitor cells cultured under serum/LPA-free (I, II) or serum/LPA-containing (III, IV) conditions with (II, IV) or without (I, III) confinement.

(A') Table summarizing cell migration characteristics for different culture conditions.

(B–D) N-Cadherin-GFP (B), Ezrin-YFP (C), and Integrin $\alpha 5$ -Integrin $\beta 1$ -BiFC-Venus (D) localization in polarized stable-bleb cells.

(E) Representative tracks of migrating stable-bleb cells.

(F) Mean-square displacement (MSD) analysis data (blue, n = 67) with fit to a persistent random walk model (gray) and compared to random (~t, orange) and directed migration (~t<sup>2</sup>, red). Black dashed line highlights random migration on timescales above the persistence time P<sub>t</sub> ~10 min (persistence length ≈ 150  $\mu$ m).

(F') Sketch of MSD time point analysis. Arrows denote variations in the direction of movement affecting persistence of cell movement.

(G) Average instantaneous stable-bleb cell migration speeds in confinement between glass-agarose or glass-glass assays for varying substrate coatings as indicated.

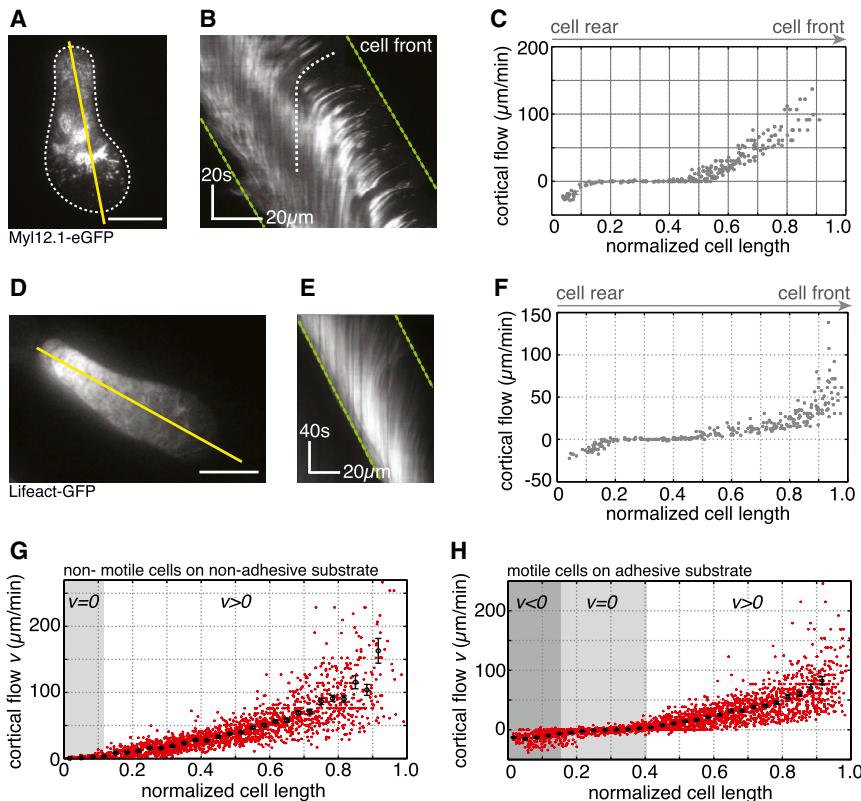
(H–J) Localization of Paxillin-GFP for stable-bleb cells on BSA- (H), PEG- (I), and Fibronectin-coated (J) glass substrates in confinement.

(K) Paxillin-GFP localization for mesodermal progenitors cultured on 2D planar substrates coated with Fibronectin. n denotes number of cells. White dashed lines in (H)–(K) outline cell contact areas. All scale bars represent 10  $\mu$ m.

See also Extended Experimental Procedures.

To test the role of specific versus unspecific adhesion, we monitored stable-bleb cell migration in various adhesive and non-adhesive environments using a modified assay that consists of two parallel glass slides coated with either Fibronectin, poly-L-lysine (PLL), BSA, or polyethyleneglycol (PEG) separated by 15- $\mu$ m beads. Surprisingly, stable-bleb cells were highly motile under all conditions with their migration efficacy increasing in non-adhesive environments (Figure 5G). To monitor the formation of adhesive foci in stable-bleb cells during migration, we further imaged Paxillin-GFP, an adaptor protein known to be involved in Integrin-mediated force transduction and focal adhesion assembly in motile cells (Turner, 2000). When stable-bleb

cells were placed between Fibronectin-coated surfaces, Paxillin-GFP localized in distinct spots at cell-substrate contacts similar to mesenchymal mesodermal cells cultured on 2D planar Fibronectin-coated substrates, while Paxillin-GFP appeared homogeneously distributed on non-adhesive BSA and PEG-coated surfaces (Figures 5H–K). This suggests that stable-bleb cells are able to generate both Integrin-dependent and independent surface attachments during migration. To further investigate the role of cell-substrate versus cell-cell adhesion, we also reduced E-Cadherin expression in stable-bleb cells by injecting antisense morpholino oligonucleotides (MOs) and found that E-Cadherin is dispensable for cell migration in vitro irrespective



**Figure 6. Frictional Coupling of Retrograde Cortical Flows Drives Stable-Bleb Migration**

(A and D) TIRFM images of (A) Myl12.1-eGFP and (D) Lifeact-GFP localization in motile stable-bleb cells. White dashed line marks the cell-to-substrate contact area. Scale bars represent 10  $\mu$ m.

(B and E) Kymograph data along the direction of movement indicated by yellow lines in (A) and (D). Green dashed lines outline cell front and rear advancement. White dashed line highlights the rearward flow of a representative fluorescent patch.

(C and F) Retrograde cortical flow velocity data in single stable-bleb cells obtained from kymograph data in (B) and (E) with  $v_{\text{Cell}} = 24.1 \mu\text{m}/\text{min}$  (C) and  $v_{\text{Cell}} = 18.2 \mu\text{m}/\text{min}$  (F).

(G and H) Average (black) and single cell local flow velocity data (red) of migrating stable-bleb cells (H) ( $n = 15$ ) with average cell speed  $v_{\text{Cell}} = 19.2 \pm 2.4 \mu\text{m}/\text{min}$  and of non-motile stable-bleb cells (G) ( $n = 30$ ) cultured in confinement under agarose on a non-adhesive PEG-coated substrate ( $v_{\text{Cell}} = 0$ ). Shaded areas denote regions with zero ( $v = 0$ , light gray) or anterograde ( $v < 0$ , dark gray) cortical flow. Cortical flows are shown in the laboratory frame of reference. All cells were cultured in the presence of 30  $\mu\text{M}$  LPA in confinement under agarose on Fibronectin-coated substrates despite data in (G). n, number of cells. Error bars represent mean  $\pm$  SEM.

See also Figure S4, Movie S5, and Extended Experimental Procedures.

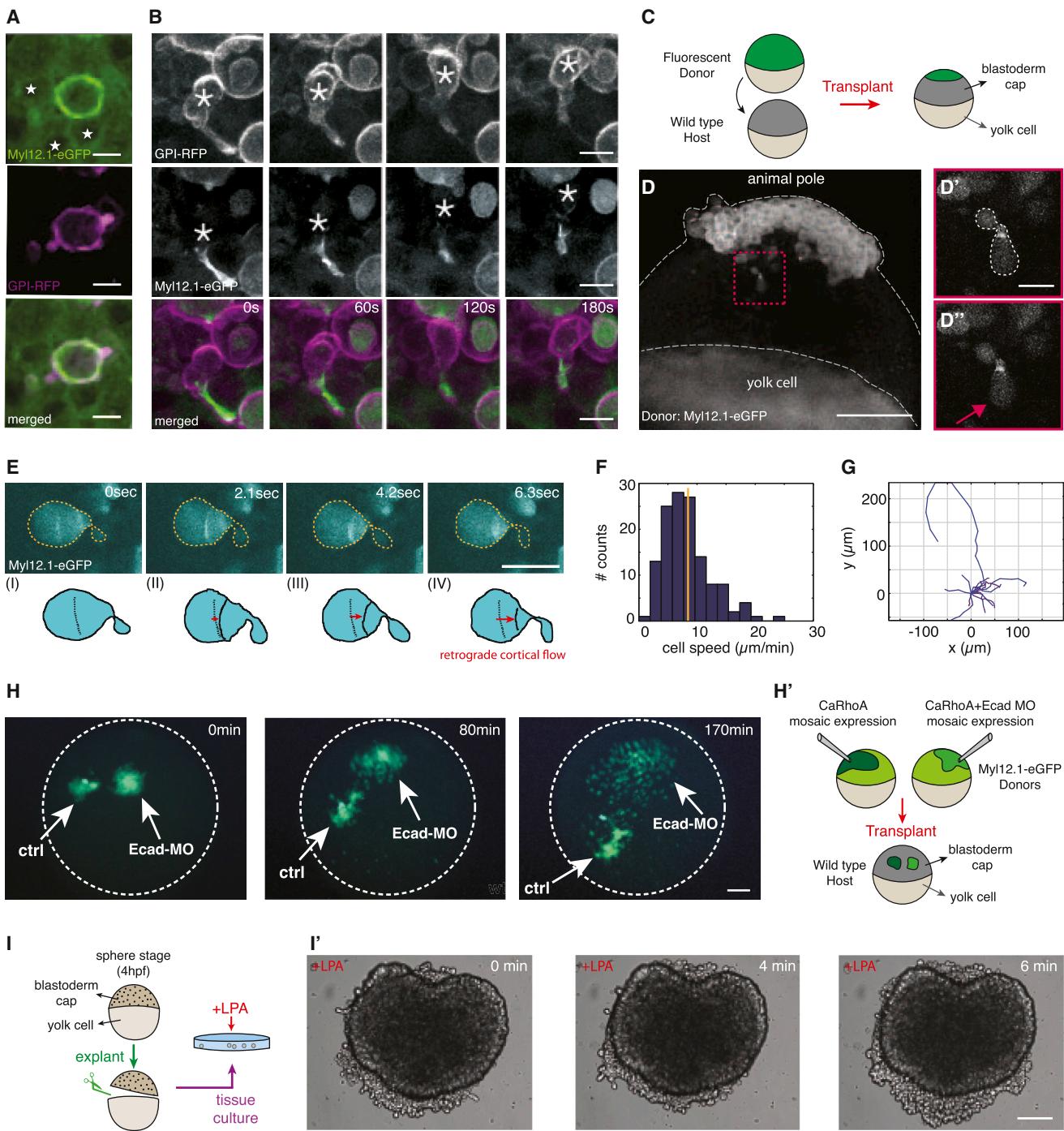
of the substrate coating (Figure S4F). Together, these data support that stable-bleb cells are able to generate both specific and unspecific traction forces during migration, highlighting their versatile migration competence in various adhesive and non-adhesive environments.

To understand this efficient and persistent motile behavior, we asked how stable-bleb cells move in 2D confined environments. Depletion of cortical elements in the cell front (Figure 4A) and the diffuse zone of actin net polymerization in the spherical front (Figure 4D), exclude a classical polymerization-driven locomotion strategy, where polymerizing actin filaments push the leading edge plasma membrane forward. However, the presence of fast retrograde cortical flows led us to hypothesize that coupling of the flowing actin network to the surrounding substrate powers stable-bleb cell migration under 2D confinement. To test this hypothesis, we performed high-resolution live cell imaging of retrograde cortical flows in stable-bleb cells in confinement under agarose. Motile stable-bleb cells on adhesive Fibronectin-coated substrates showed a rapid decay of flow velocities toward the cell rear with a pronounced zone of zero flow speed (Figures 6A–6F and 6H; Movie S5). In contrast, non-motile stable-bleb cells that were placed on passivated PEG-coated substrates revealed considerably different retrograde flow profiles with similar maximal flow velocities to motile cells in the cell front, but a slower decay of flow speeds to the cell rear (Figures 6G and S3D). Using a 1D fluid description of the cortex, we estimated from those cortical flow profiles a higher frictional coefficient  $\xi$  on adhesive in comparison to non-adhesive surfaces ( $\xi_{\text{adhesive}} \approx 2.8 \times 10^8$

$\text{Pa} \times \text{s/m}$  versus  $\xi_{\text{non-adhesive}} \approx 1 \times 10^8 \text{ Pa} \times \text{s/m}$ ; Figures S4A–S4C). Furthermore, osmotic shock experiments revealed that cortical flows are progressively decreased for lower levels of intracellular pressure and that these changes in cortical flow velocity strongly correlate with cell migration speed (Figures S4D and S4E). These findings support the notion that coupling of the cortex to the surrounding substrate leads to a reduction of cortical flow in motile cells, providing the necessary forward traction force required for cell movement.

#### Evidence for Stable-Bleb Cell Migration In Vivo

To study the motile switch and migration characteristics of stable-bleb cells *in vivo*, we first aimed to ectopically transform cells in the early zebrafish blastula by elevating their cortical contractility. To this end, we expressed a constitutively active version of RhoA (caRhoA) (Takesono et al., 2012) in single progenitor cells and performed two-photon imaging of their morphology and migration properties in the gastrulating embryo. CaRhoA-expressing cells *in vivo* showed high blebbing rates and pronounced cortical myosin accumulations similar to LPA-stimulated progenitor cells *in vitro* (Figure 7A). Moreover, caRhoA-expressing cells frequently transformed into motile cells closely resembling stable-bleb cells *in vitro* displaying a pear-shape, pronounced cortical actin and myosin II gradients and rearward cortical flows of similar magnitude as found *in vitro* ( $v_{\text{cortex}} \sim 30 \mu\text{m}/\text{min}$ ; Figures 7B and S5A–S5C; Movie S6). To further test whether caRhoA expressing progenitor cells *in vivo* indeed resemble stable-bleb cells *in vitro*, we transferred them into



**Figure 7. Identification and Characterization of Stable-Bleb Cells In Vivo**

(A) Myl12.1-eGFP (green) and GPI-RFP (magenta) localization in non-polarized embryonic cells observed in a *Tg(actβ1:myl12.1eGFP)* transgenic embryo at sphere stage expressing caRhoA and GPI-RFP in a mosaic pattern. Asterisks denote cell nuclei in the vicinity of a caRhoA expressing cell highlighting cells with reduced cortical accumulation of myosin II compared to caRhoA expressing cells. Scale bars, 10 μm.

(B) Fluorescence time-lapse series of a stable-bleb like cell in a *Tg(actβ1:myl12.1eGFP)* transgenic embryo at sphere stage expressing caRhoA and GPI-RFP in a mosaic pattern. Scale bars represent 10 μm.

(C) Illustration of the transplantation experiment in (D).

(D) Transplantation of cells from a *Tg(actβ1:myl12.1eGFP)* transgenic donor to a wild-type host embryo at sphere stage. A stable-bleb-like cell emerging from the transplanted tissue is highlighted in the red box. Scale bar represents 100 μm.

(legend continued on next page)

culture and found that they exhibited a morphology and migration behavior indistinguishable from LPA-induced stable-bleb cells *in vitro* (Figures S4G–S4J). This suggests that upregulation of cortical contractility through expression of caRhoA is sufficient to transform progenitor cells into motile stable-bleb cells *in vivo*.

To unravel the role of cell-cell adhesion for adhesive force coupling during stable-bleb cell motility *in vivo*, we downregulated E-Cadherin expression in caRhoA expressing embryos by injecting specific *e-Cadherin* MOs. Upon transplantation of those cells into a wild-type host we found that E-Cadherin was dispensable for stable-bleb migration *in vivo* and even led to an increase of cell dispersal compared to caRhoA expressing cells with normal levels of E-Cadherin (Figures 7H and S5D–S5F).

We next sought to address if stable-bleb cells also emerge in the endogenous context of embryonic development. Time-lapse movies of cell movements within the gastrulating embryo occasionally revealed single fast migrating amoeboid cells reminiscent of stable-bleb cells (see Movie S7). However, since it was impossible to predict where and when in the gastrulating embryo these cells appear, we reasoned that locally increasing cortical contractility within the gastrulating embryo might trigger stable-bleb cell transformation. Wounding sites have previously been shown to display elevated levels of actomyosin contractility (Sonnemann and Bement, 2011). To test if stable-bleb cells emerge at local wounding sites within the gastrulating embryo, we transplanted fluorescently labeled donor cells into an unlabeled host embryo at sphere stage (4 hr post fertilization [hpf]) (Figures 7C and 7D). Using *Tg(actβ1:myl12.1eGFP)* donor embryos, elevated cortical accumulations of myosin II were detected around the wounding area (Movie S8) and, interestingly, labeled donor cells stochastically transformed into stable-bleb like cells with a highly polarized cortical architecture and fast retrograde cortical flows similar to their morphology *in vitro* (Figure 7E; Movie S8). Transformed cells migrated away from the region of transplantation with speed ( $v = 8.5 \pm 0.3 \mu\text{m}/\text{min}$ ) and directionality much higher than any other cells yet described in the zebrafish embryo (Figures 7F and 7G). Together, this indicates that stable-bleb motility can be induced at places of high contractility within the gastrulating embryo, suggesting a mechanism for rapid cell extrusion from highly contractile regions within the embryo. To further test this assumption, we prepared tissue explants from early blastula stage embryos and cultured

them in LPA containing medium. We found that marginal cells in tissue explants not only rapidly transformed into stable-bleb cells, as expected for LPA-treated progenitors, but also frequently extruded from the explant surface (Figure 7I). This suggests that stable-bleb cells can overcome the surface tension of those tissue explants and freely disperse into the environment, supporting our notion of stable-bleb cell transformation as an efficient cell extrusion mechanism from contractile tissues.

## DISCUSSION

Here, we describe a stochastic amoeboid motility switch of early zebrafish progenitor cells into stable-bleb cells. Cell morphology and migration characteristics of stable-bleb cells show unique features in comparison to other motile cell types described before. The invariant roundish cortex-depleted front end of stable-bleb cells is clearly different from actin rich protrusions, such as lamellipodia and filopodia typically found in mesenchymal cells and also distinct from contractility-driven bleb protrusions that periodically appear in blebbing cells. We show that stable polarization of transformed progenitor cells is maintained by a positive feedback loop between cortical density gradients and continuous cortical flows, yielding a robust steady-state configuration with a highly polarized cortical architecture. Interestingly, we found that increasing cortical contractility is sufficient to switch progenitor cells from a blebbing or mesenchymal mode to a stably polarized state above a critical threshold level of contractility, and that this transformation is reversible upon lowering myosin II activity below a critical threshold level. This indicates that the level of cortical contractility determines whether cells migrate by forming either transient protrusions, such as blebs and lamellipodia, or one stably polarized front thereby increasing their migration persistence and efficacy. Supported by theoretical modeling of progenitor transformation into stable-bleb cells, it further suggests that this process does not require an elaborate genetic program, but rather occurs as a simple mechanical switch triggered by increased cortical contractility. This generic polarization module can likely be activated via diverse cellular pathways such as microtubule disassembly-dependent activation of the Rho/ROCK pathway (Niggli, 2003) and could thus be a potentially effective polarization mechanism also functional in other cell types. Notably, findings in various culture cell lines suggest that high spatial confinement combined with low substrate adhesion is sufficient to trigger an

(D' and D') Magnified views on transplanted stable-bleb-like cells observed in (D). Dashed white line marks cell borders. Red arrow indicates the spherical protrusion front. Scale bar represents 20  $\mu\text{m}$ .

(E) Myl12.1-eGFP localization and cortical flow in a stable-bleb cell obtained from transplantation of *Tg(actβ1:myl12.1eGFP)* transgenic donor cells into a wild-type host. Red arrows point to the direction of cortical flow. Dashed yellow lines outline cell borders. Scale bar represents 25  $\mu\text{m}$ .

(F and G) Representative tracks (F) and instantaneous migration speeds (G) of stable-bleb cells *in vivo* ( $n = 15$  cells) with average velocity ( $\langle v_{\text{cell}} \rangle = 8.5 \pm 0.3 \mu\text{m}/\text{min}$ ; yellow line).

(H) Time-lapse fluorescence images of a cell transplantation between Myl12.1-eGFP donor embryos expressing caRhoA (Ctrl) or caRhoA+E-Cadherin morpholino (Ecad-MO) and a wild-type host at dome stage. Dashed white lines outline the embryo margin. Scale bar represents 100  $\mu\text{m}$ .

(H') Schematic illustration of the transplantation experiment in (H).

(I) Sketch of tissue explant preparation and culturing.

(I') BF time-lapse images of a blastoderm tissue explant obtained from wild-type embryos at sphere stage and cultured with 30  $\mu\text{M}$  LPA. Scale bar represents 100  $\mu\text{m}$ .

See also Figure S5, Movies S6, S7, and S8, and Extended Experimental Procedures.

amoeboid cell transformation strikingly similar to stable-bleb cells (see Liu et al., 2015). Together, these data suggest that the regulation of cortical contractility might serve as a universal switch for transforming cells into an amoeboid cell migration mode characterized by fast and highly directional migration in 3D confinement.

Spatially confined environments promote stable-bleb cell migration by providing compressive forces that increase frictional coupling required for efficient locomotion. Moreover, retrograde cortical flows are sufficient to drive stable-bleb cell migration in both adhesive and non-adhesive confined environments. Similar flow-friction mechanisms have recently been proposed to drive locomotion of various cell types in confinement (Hawkins et al., 2009; Liu et al., 2015), suggesting that force-coupling via retrograde cortical flows might constitute a general mechanism of cell movement especially in low adhesive 3D environments.

During gastrulation cell motility needs to be precisely controlled and spatio-temporally regulated. Recent studies have pointed at important functions of small Rho GTPases in cell migration during zebrafish gastrulation (Matthews et al., 2008; Kardash et al., 2010). Our observation that LPA, a RhoA-activating signal ubiquitously expressed during zebrafish gastrulation (Lee et al., 2008; Yukiura et al., 2011), can transform progenitor cells into stable-bleb cells, reveals a yet unrecognized function of RhoA in the control of cell migration during gastrulation and points at LPA as one potential regulator determining stable-bleb cell induction during gastrulation. However, given the large number of known signals with RhoA activating capacity, a systematic functional analysis of these signals will be needed to dissect the upstream regulation of stable-bleb cell transformation during gastrulation.

As to the function of stable-bleb cells in gastrulation, we observed stable-bleb cell transformation at areas of high cell contractility, such as wounding sites, followed by rapid dispersal of those cells. This suggests that stable-bleb transformation serves as a mechanism for cell extrusion from areas of increased contractility in the embryo. This might help in equilibrating contractility-driven cell density variations, but might also be used as a mean of long-range communication where extruding stable-bleb cells rapidly transmit signals from areas of high contractility to the remainder of the embryo. Our observation of stable-bleb cell extrusion from tissue explants along with observations of a similar migration phenotype in various cancer cell lines (see Liu et al., 2015) also points at the intriguing possibility that stable-bleb cells might be involved in other contexts such as cancer metastasis where cancer cells disseminate from the primary tumor into unaffected tissues. Thus, stable-bleb cell transformation identifies a yet unrecognized mechanism for rapid and long-range cell dispersal in vertebrates.

## EXPERIMENTAL PROCEDURES

### Zebrafish Maintenance

Zebrafish (*Danio rerio*) were maintained as previously described (Westerfield, 2007). Embryos were kept in E3 medium at 25°C–31°C prior to experiments and staged based on morphological criteria (Kimmel et al., 1995) and hours postfertilization (hpf). Wild-type embryos were obtained from the Tup Longfin (TL) background unless indicated otherwise. See Extended Experimental Procedures for transgenic fish lines and mRNA/morpholino injections.

### Cell Culture

To culture primary germ layer progenitor cells, embryos were manually dechorionated in E3 buffer at sphere stage (4 hpf) or 75% epiboly stage (8 hpf). Five to ten embryos were transferred to DMEM-F12 culture medium (Sigma) and mechanically dissociated by manual tapping followed by centrifugation at 200 × g for 2 min. To induce stable-bleb polarization, 100 μM LPA was added during centrifugation unless indicated otherwise. See Extended Experimental Procedures for more details.

### Reagents and Inhibitor Treatments

Fetal BSA (GIBCO) and 1-Oleoyl lysophosphatidic acid (LPA, Tocris Bioscience) were used at the indicated concentrations. Fibroblast growth factor mFGF-8b and rhFGF-3 (R&D Bioscience), platelet-derived growth factor PDGF-AA (PeproTech), and epidermal growth factor hEGF (Sigma) were used at a final concentration of 100 ng/μl. Pharmacological inhibitors were used at the following concentrations: 10 μM active (–) or inactive Blebbistatin (+) (Tocris Bioscience), 100 nM Latrunculin-A (Sigma), 100 μM Y-27632 (Tocris Bioscience), 500 nM Jasplakinolide (Tocris Bioscience), 20 μM ML-141 (Sigma), and 30 μM L-294002 (Sigma). The percentage of polarized cells was measured after 15 min (Blebbistatin, LatrunculinA, and Jasplakinolide) and 30–60 min (Y-27632, ML-141, Y-294002) of exposure.

### Data Statistics

Data are presented as mean ± SEM unless indicated otherwise. Statistical tests were performed by using the Wilcoxon rank sum test in MATLAB.

## SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, five figures, and eight movies and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2015.01.008>.

## AUTHOR CONTRIBUTIONS

V.R. and S.W. designed research, performed experiments, and analyzed the data. A.C.-J. and R.V. developed the theoretical model. M.S., H.M., and K.S. helped with *in vivo* experiments. V.B. generated the mezzo-GFP fishline and performed micropipette experiments. M.R.-M. and M.S. supported the work of V.R. and S.W. V.R. and C.-P.H. supervised the project and wrote the manuscript.

## ACKNOWLEDGMENTS

We would like to thank R. Hausschild and E. Papusheva for technical assistance and the service facilities at the IST Austria for continuous support. The caRhoA plasmid was a kind gift of T. Kudoh and A. Takesono. We thank M. Piel and E. Paluch for exchanging unpublished data. This work was supported by grants from the Agence Nationale de la Recherche (ANR) to R.V. (ANR-09-PIRI-0027), the Human Frontiers in Science Program to M.S., and the Austrian Science Foundation (FWF): (T560-B17) and (I 812-B12) to V.R. and C.-P.H.

Received: April 15, 2014

Revised: October 15, 2014

Accepted: December 12, 2014

Published: February 12, 2015

## REFERENCES

- Behrndt, M., Saibreux, G., Campinho, P., Hauschild, R., Oswald, F., Roensch, J., Grill, S.W., and Heisenberg, C.-P. (2012). Forces driving epithelial spreading in zebrafish gastrulation. *Science* 338, 257–260.
- Blaser, H., Reichman-Fried, M., Castanon, I., Dumstrei, K., Marlow, F.L., Kawakami, K., Solnica-Krezel, L., Heisenberg, C.-P., and Raz, E. (2006). Migration of zebrafish primordial germ cells: a role for myosin contraction and cytoplasmic flow. *Dev. Cell* 11, 613–627.

Bois, J.S., Jülicher, F., and Grill, S.W. (2011). Pattern formation in active fluids. *Phys. Rev. Lett.* **106**, 028103.

Callan-Jones, A.C., and Voituriez, R. (2013). Active gel model of amoeboid cell motility. *New J. Phys.* **15**, 025022.

Carvalho, K., Tsai, F.C., Lees, E., Voituriez, R., Koenderink, G.H., and Sykes, C. (2013). Cell-sized liposomes reveal how actomyosin cortical tension drives shape change. *Proc. Natl. Acad. Sci. USA* **110**, 16456–16461.

Diz-Muñoz, A., Krieg, M., Bergert, M., Ibarlucea-Benitez, I., Muller, D.J., Paluch, E., and Heisenberg, C.-P. (2010). Control of directed cell migration in vivo by membrane-to-cortex attachment. *PLoS Biol.* **8**, e1000544.

Friedl, P., and Alexander, S. (2011). Cancer invasion and the microenvironment: plasticity and reciprocity. *Cell* **147**, 992–1009.

Hawkins, R.J., Piel, M., Faure-Andre, G., Lennon-Dumenil, A.M., Joanny, J.F., Prost, J., and Voituriez, R. (2009). Pushing off the walls: a mechanism of cell motility in confinement. *Phys. Rev. Lett.* **102**, 058103.

Hawkins, R.J., Poincloux, R., Bénichou, O., Piel, M., Chavrier, P., and Voituriez, R. (2011). Spontaneous contractility-mediated cortical flow generates cell migration in three-dimensional environments. *Biophys. J.* **101**, 1041–1045.

Kardash, E., Reichman-Fried, M., Maître, J.-L., Boldajipour, B., Papusheva, E., Messerschmidt, E.-M., Heisenberg, C.-P., and Raz, E. (2010). A role for Rho GTPases and cell-cell adhesion in single-cell motility in vivo. *Nat. Cell Biol.* **12**, 47–53, 1–11.

Kimmel, C.B., Ballard, W.W., Kimmel, S.R., Ullmann, B., and Schilling, T.F. (1995). Stages of embryonic development of the zebrafish. *Dev. Dyn.* **203**, 253–310.

Lämmermann, T., and Sixt, M. (2009). Mechanical modes of ‘amoeboid’ cell migration. *Curr. Opin. Cell Biol.* **21**, 636–644.

Lee, S.-J., Chan, T.-H., Chen, T.-C., Liao, B.-K., Hwang, P.-P., and Lee, H. (2008). LPA1 is essential for lymphatic vessel development in zebrafish. *FASEB J.* **22**, 3706–3715.

Liu, Y.-J., Le Berre, M., Lautenschlaeger, F., Maiuri, P., Callan-Jones, A., Heuzé, M., Takaki, T., Voituriez, R., and Piel, M. (2015). Confinement and low adhesion induce fast amoeboid migration of slow mesenchymal cells. *Cell* **160**, this issue, 660–673.

Matthews, H.K., Marchant, L., Carmona-Fontaine, C., Kuriyama, S., Larraín, J., Holt, M.R., Parsons, M., and Mayor, R. (2008). Directional migration of neural crest cells in vivo is regulated by Syndecan-4/Rac1 and non-canonical Wnt signaling/RhoA. *Development* **135**, 1771–1780.

Mayer, M., Depken, M., Bois, J.S., Jülicher, F., and Grill, S.W. (2010). Anisotropies in cortical tension reveal the physical basis of polarizing cortical flows. *Nature* **467**, 617–621.

Mills, G.B., and Moolenaar, W.H. (2003). The emerging role of lysophosphatidic acid in cancer. *Nat. Rev. Cancer* **3**, 582–591.

Montero, J.-A., Kilian, B., Chan, J., Bayliss, P.E., and Heisenberg, C.-P. (2003). Phosphoinositide 3-kinase is required for process outgrowth and cell polarization of gastrulating mesendodermal cells. *Curr. Biol.* **13**, 1279–1289.

Niggli, V. (2003). Microtubule-disruption-induced and chemotactic-peptide-induced migration of human neutrophils: implications for differential sets of signalling pathways. *J. Cell Sci.* **116**, 813–822.

Poincloux, R., Collin, O., Lizárraga, F., Romao, M., Debray, M., Piel, M., and Chavrier, P. (2011). Contractility of the cell rear drives invasion of breast tumor cells in 3D Matrigel. *Proc. Natl. Acad. Sci. USA* **108**, 1943–1948.

Row, R.H., Maître, J.-L., Martin, B.L., Stockinger, P., Heisenberg, C.-P., and Kimelman, D. (2011). Completion of the epithelial to mesenchymal transition in zebrafish mesoderm requires Spadetail. *Dev. Biol.* **354**, 102–110.

Salbreux, G., Prost, J., and Joanny, J.F. (2009). Hydrodynamics of cellular cortical flows and the formation of contractile rings. *Phys. Rev. Lett.* **103**, 058102.

Shih, W., and Yamada, S. (2010). Myosin IIA dependent retrograde flow drives 3D cell migration. *Biophys. J.* **98**, L29–L31.

Sonnemann, K.J., and Bement, W.M. (2011). Wound repair: toward understanding and integration of single-cell and multicellular wound responses. *Annu. Rev. Cell Dev. Biol.* **27**, 237–263.

Takesono, A., Moger, J., Farooq, S., Cartwright, E., Dawid, I.B., Wilson, S.W., and Kudoh, T. (2012). Solute carrier family 3 member 2 (Slc3a2) controls yolk syncytial layer (YSL) formation by regulating microtubule networks in the zebrafish embryo. *Proc. Natl. Acad. Sci. USA* **109**, 3371–3376.

Tinevez, J.-Y., Schulze, U., Salbreux, G., Roensch, J., Joanny, J.-F., and Paluch, E. (2009). Role of cortical tension in bleb growth. *Proc. Natl. Acad. Sci. USA* **106**, 18581–18586.

Tozluoğlu, M., Tournier, A.L., Jenkins, R.P., Hooper, S., Bates, P.A., and Sahai, E. (2013). Matrix geometry determines optimal cancer cell migration strategy and modulates response to interventions. *Nat. Cell Biol.* **15**, 751–762.

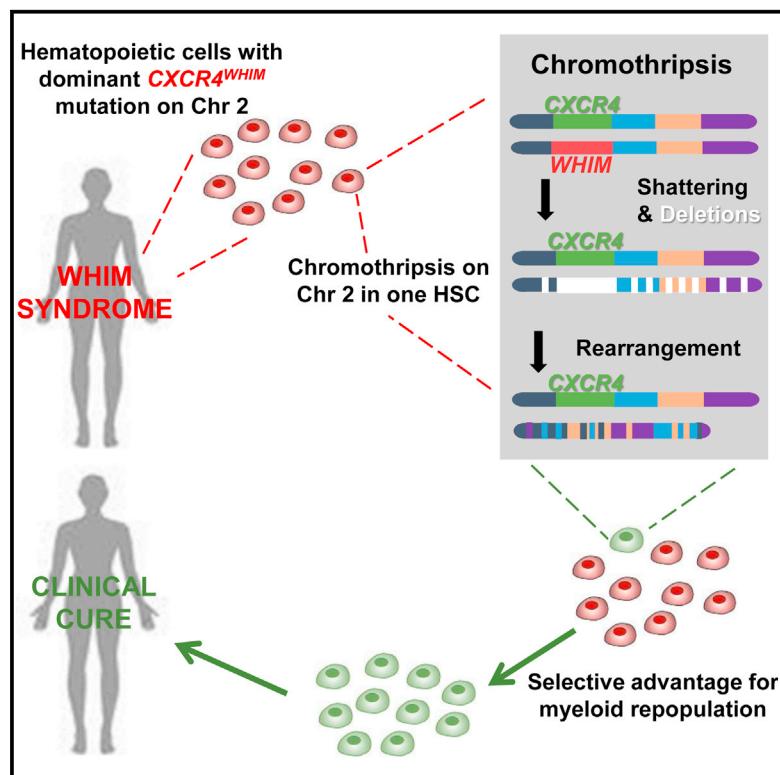
Turner, C.E. (2000). Paxillin and focal adhesion signalling. *Nat. Cell Biol.* **2**, E231–E236.

Westerfield, M. (2007). The Zebrafish Book. A Guide for the Laboratory Use of Zebrafish (*Danio rerio*), Fifth Edition (Eugene: University of Oregon Press).

Yukiura, H., Hama, K., Nakanaga, K., Tanaka, M., Asaoka, Y., Okudaira, S., Arima, N., Inoue, A., Hashimoto, T., Arai, H., et al. (2011). Autotaxin regulates vascular development via multiple lysophosphatidic acid (LPA) receptors in zebrafish. *J. Biol. Chem.* **286**, 43972–43983.

# Chromothriptic Cure of WHIM Syndrome

## Graphical Abstract



## Authors

David H. McDermott, Ji-Liang Gao, ..., Harry L. Malech, Philip M. Murphy

## Correspondence

pmm@nih.gov

## In Brief

WHIM syndrome is an inherited immunodeficiency caused by overactivity of CXCR4, a receptor controlling production and distribution of leukocytes in bone marrow and blood. We identified a WHIM patient cured by chromothripsy (chromosome shattering) that fortuitously deleted the abnormal copy of the CXCR4 gene in a single hematopoietic stem cell, which then took over the bone marrow and restored normal immune function. This experiment of nature suggests that partial CXCR4 inactivation might enhance engraftment of bone marrow in patients requiring transplantation.

## Highlights

- CXCR4 haploinsufficiency may promote HSC engraftment
- Chromothriptic deletions result in functional cure of WHIM syndrome
- Clinical symptoms in WHIM syndrome are dependent on myeloid deficits
- Myeloid-derived cells are critical for control of HPV infection

# Chromothriptic Cure of WHIM Syndrome

David H. McDermott,<sup>1,14</sup> Ji-Liang Gao,<sup>1,14</sup> Qian Liu,<sup>1</sup> Marie Siwicki,<sup>1</sup> Craig Martens,<sup>2</sup> Paejonette Jacobs,<sup>1</sup> Daniel Velez,<sup>1</sup> Erin Yim,<sup>1</sup> Christine R. Bryke,<sup>4,12</sup> Nancy Hsu,<sup>4,12</sup> Zunyan Dai,<sup>4,13</sup> Martha M. Marquesen,<sup>5</sup> Elina Stregevsky,<sup>3</sup> Nana Kwatema,<sup>5</sup> Narda Theobald,<sup>5</sup> Debra A. Long Priel,<sup>6</sup> Stefania Pittaluga,<sup>7</sup> Mark A. Raffeld,<sup>7</sup> Katherine R. Calvo,<sup>8</sup> Irina Maric,<sup>8</sup> Ronan Desmond,<sup>9,10</sup> Kevin L. Holmes,<sup>3</sup> Douglas B. Kuhns,<sup>6</sup> Karl Balabanian,<sup>11</sup> Françoise Bachelerie,<sup>11</sup> Stephen F. Porcella,<sup>2</sup> Harry L. Malech,<sup>5</sup> and Philip M. Murphy<sup>1,\*</sup>

<sup>1</sup>Laboratory of Molecular Immunology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892, USA

<sup>2</sup>Research Technologies Branch, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Hamilton, MT 59840, USA

<sup>3</sup>Research Technologies Branch, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892, USA

<sup>4</sup>Quest Diagnostics, Chantilly, VA 20151, USA

<sup>5</sup>Laboratory of Host Defenses, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892, USA

<sup>6</sup>Clinical Services Program, Leidos Biomedical Research, Frederick National Laboratory for Cancer Research, Frederick, MD 21701, USA

<sup>7</sup>Laboratory of Pathology, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA

<sup>8</sup>Division of Laboratory Medicine, Clinical Center, National Institutes of Health, Bethesda, MD 20892, USA

<sup>9</sup>Hematology Branch, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD 20892, USA

<sup>10</sup>Department of Haematology, Tallaght Hospital, Dublin 24, Ireland

<sup>11</sup>INSERM UMR- S996, Laboratoire d'Excellence in Research on Medication and Innovative Therapeutics, Université Paris-Sud, 92140 Clamart, France

<sup>12</sup>Department of Cytogenetics, Beth Israel Deaconess Medical Center, Boston, MA 02215, USA

<sup>13</sup>Department of Human Genetics, Emory University School of Medicine, Decatur, GA 30030, USA

<sup>14</sup>Co-first author

\*Correspondence: pmm@nih.gov

<http://dx.doi.org/10.1016/j.cell.2015.01.014>

## SUMMARY

**Chromothripsy is a catastrophic cellular event recently described in cancer in which chromosomes undergo massive deletion and rearrangement. Here, we report a case in which chromothripsy spontaneously cured a patient with WHIM syndrome, an autosomal dominant combined immunodeficiency disease caused by gain-of-function mutation of the chemokine receptor CXCR4. In this patient, deletion of the disease allele, CXCR4<sup>R334X</sup>, as well as 163 other genes from one copy of chromosome 2 occurred in a hematopoietic stem cell (HSC) that repopulated the myeloid but not the lymphoid lineage. In competitive mouse bone marrow (BM) transplantation experiments, Cxcr4 haploinsufficiency was sufficient to confer a strong long-term engraftment advantage of donor BM over BM from either wild-type or WHIM syndrome model mice, suggesting a potential mechanism for the patient's cure. Our findings suggest that partial inactivation of CXCR4 may have general utility as a strategy to promote HSC engraftment in transplantation.**

## INTRODUCTION

WHIM syndrome is an autosomal dominant combined primary immunodeficiency disease caused by mutations in the chemokine receptor CXCR4 (Hernandez et al., 2003). The term "WHIM" is an acronym for the main manifestations of the dis-

ease: warts, hypogammaglobulinemia, recurrent infections, and myelokathexis; myelokathexis refers to impaired egress of mature neutrophils from bone marrow causing neutropenia (Wetzler et al., 1990; Zuelzer, 1964). Most patients with WHIM syndrome are actually panleukopenic, with severely reduced peripheral blood B cells but less severe reductions in peripheral blood T cells and monocytes (McDermott et al., 2011). The signature pathogen in WHIM syndrome is human papillomavirus (HPV), which causes warts that cannot be controlled with standard medical treatment and may progress to cancer (Al Ustwani et al., 2014; Beaussant Cohen et al., 2012; Dotta et al., 2011; Kawai and Malech, 2009). Recurrent bacterial infections also occur, mainly in the sinopulmonary tract, oral cavity, ear, skin, and soft tissue, where chronic complications may arise, especially bronchiectasis and hearing loss. Prophylactic antibiotics, IVIg and G-CSF are often used to reduce the incidence of infections; however, their precise efficacy has not been established (Al Ustwani et al., 2014; Beaussant Cohen et al., 2012; Dotta et al., 2011; Kawai and Malech, 2009). In contrast, safety and preliminary evidence of clinical efficacy has recently been reported from Phase I studies of the specific CXCR4 antagonist plerixafor (Mozobil, AMD3100) (Dale et al., 2011; McDermott et al., 2011, 2014). Spontaneous remission or cure of WHIM syndrome has not been previously reported.

WHIM mutations of CXCR4 increase signaling because they disrupt negative regulatory elements in the carboxy-terminus, thereby exaggerating the normal hematopoietic functions of the receptor (Haribabu et al., 1997; Signoret et al., 1998; Venkatesan et al., 2003). CXCR4 is normally expressed by most leukocytes and has one ligand, CXCL12 (Bachelerie et al., 2014; Bleul et al., 1996), which is constitutively expressed at high levels by stromal cells in the bone marrow and normally mediates HSC

retention in bone marrow niches (Broxmeyer et al., 2003b; Broxmeyer et al., 2005; Dar et al., 2006; Sugiyama et al., 2006; Zou et al., 1998). In addition, CXCR4 signaling promotes hematopoietic stem cell (HSC) quiescence, homing to bone marrow from blood and differentiation into committed myeloid progenitors (Broxmeyer et al., 2003a, 2003b; Kawai et al., 2007; Nie et al., 2008; Sugiyama et al., 2006).

Chromothripsis refers to multiple clustered genetic rearrangements and deletions affecting one or a few chromosomes (Stephens et al., 2011). The abnormalities are thought to occur all at once in a single cell, which then presumably either dies or acquires a growth advantage, depending on the genes affected (Stephens et al., 2011). Accordingly, chromothripsis was first identified by whole genome sequencing of cancer cell lines and has been reported to affect ~2% of all cancers (Jones and Jallepalli, 2012), as well as one patient with a severe congenital cognitive syndrome (Kloosterman et al., 2011). Criteria for chromothripsis include (1) clustering of breakpoints in limited areas of one or several chromosomes with large intervening regions of normal sequence, (2) copy number states that suddenly oscillate between areas of normal heterozygosity and loss of heterozygosity, and (3) rearrangements affecting a single haplotype with multiple fragments rearranged in random orientation and order (Korbel and Campbell, 2013). Here, we describe chromothriptic deletions of one copy of chromosome 2, including deletion of the disease allele  $CXCR4^{R334X}$ , in a patient with WHIM syndrome that resulted in cure of the disease.

## RESULTS

### Cure of a Patient with WHIM Syndrome

The index patient, designated WHIM-09, is a white female who presented at age 58 to the NIH requesting evaluation for herself and two of her three daughters, designated WHIM-10 (age 21) and WHIM-11 (age 23) (Figure 1A). Both daughters had a history of recurrent infections since early childhood, multiple cutaneous warts, panleukopenia, and hypogammaglobulinemia, and therefore fulfilled all the clinical criteria for WHIM syndrome. In contrast, WHIM-09 reported that from childhood through age 38 she had had many serious infections, often requiring hospitalization, but then none in the 20 subsequent years, and that she had had confluent warts on her hands that spontaneously resolved also in her 30s (Figure 1B). Moreover, we found that at the time of presentation, WHIM-09 was not neutropenic, but instead had a mild leukocytosis, including an absolute neutrophil count (ANC) and absolute monocyte count (AMC) that were ~2-fold greater than the upper limit of normal; in contrast, the absolute lymphocyte count (ALC) was within normal limits. The past medical history revealed that WHIM-09 was in fact the first patient ever described with myelokathexis, the key hematopathologic feature in WHIM syndrome, reported in two articles published in *The New England Journal of Medicine* (NEJM) in 1964 (Krill et al., 1964; Zuelzer, 1964). She underwent a therapeutic splenectomy at age 9 for the possibility of autoimmune neutropenia, which was ineffective. There is no evidence that her parents or siblings had WHIM syndrome. Thus, the history and clinical evidence were compatible with a WHIM mutation occurring de novo in patient WHIM-09, autosomal dominant transmission

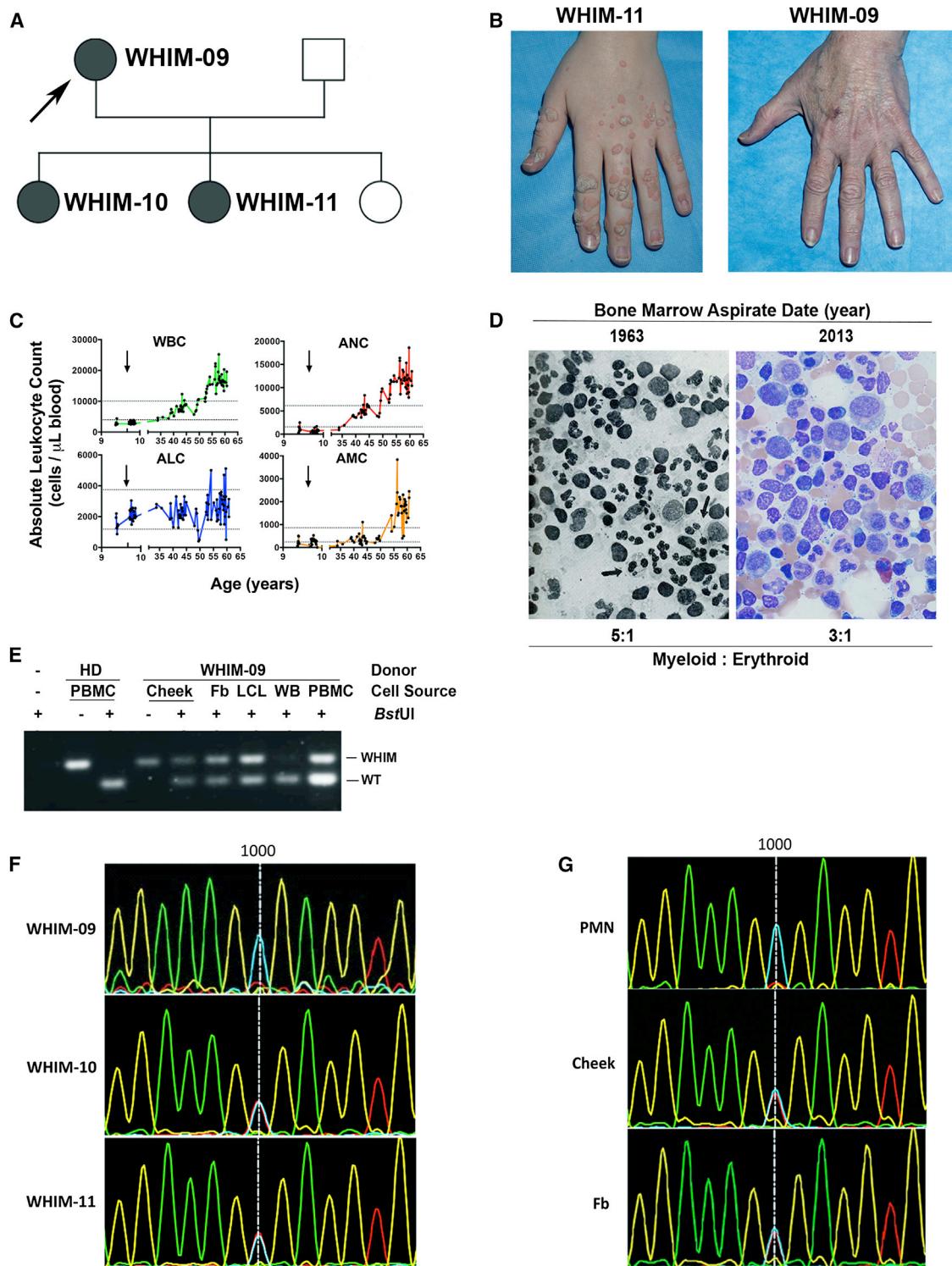
to two of her three daughters, and spontaneous and durable complete clinical remission in WHIM-09 in her fourth decade of life (Figure 1A).

To evaluate potential mechanisms for clinical remission, we first graphed all available white blood cell counts for WHIM-09, including those previously published in the NEJM (Figure 1C). Consistent with the clinical history, this revealed severe neutropenia at least from age 9 that was unaffected by splenectomy but that began to resolve spontaneously early in the fourth decade of life, rising slowly over time to a new and stable baseline slightly above the upper limit of normal. The AMC followed the same time course, whereas, interestingly, the ALC did not, starting in the normal range for healthy individuals as a child then increasing inconsistently and only slightly as an adult. Nevertheless, when lymphocyte subsets were examined in detail, all B cell subsets and both naive CD4<sup>+</sup> and CD8<sup>+</sup> T cell subsets were below the lower limit of normal (Table 1), as they were in both daughters and most other patients reported with WHIM syndrome. Consistent with this, WHIM-09 was slightly hypogammaglobulinemic at the time of presentation to NIH with IgG = 535 mg/dl (normal range, 642–1,730 mg/dl). In contrast, memory CD4<sup>+</sup> and CD8<sup>+</sup> T cell subsets were elevated in WHIM-09, but deficient in both daughters. Unfortunately, WHIM-09's archival lymphocyte subset values from the years when she fulfilled the clinical criteria for WHIM syndrome were not available.

We considered myeloid leukemia or a possible pre-leukemic state as a cause of her mild leukocytosis; however, the patient was clinically well over this ~20-year time span when her neutrophils and monocytes were increasing, and her blood smear and bone marrow histopathology at the NIH were not consistent with these diagnoses (Figure 1D); moreover, specific testing for B and T cell clonality as well as for *BCR-ABL* and *JAK* mutations were negative (see Extended Experimental Procedures for details). Consistent with her apparent ~20-year complete remission of WHIM syndrome by clinical criteria, her bone marrow did not present the characteristic features of the disease (increased myeloid:erythroid ratio, neutrophil vacuolization, eyeglass nuclei in neutrophils), which were, however, all present in her bone marrow histopathology reported in the NEJM in 1964, shown again here for comparison, with permission, in Figure 1D (Zuelzer, 1964). Since the patient reported she had undergone several prior surgeries and blood transfusions, we tested her blood for evidence of allogeneic chimerism and found none (data not shown). Thus, although the patient appeared to be clinically cured, she was hematologically mosaic, with sustained spontaneous correction of neutropenia, monocytopenia, and myelokathexis, and continued deficiency of B and naive T lymphocytes in the blood.

### Patient WHIM-09 Is a Genetic Mosaic for $CXCR4^{R334X}$ , the Most Common Mutation Causing WHIM Syndrome

We next genotyped whole blood DNA from family members using an established PCR-BstUI restriction fragment length polymorphism (PCR-RFLP) assay for the most common mutation in WHIM syndrome,  $CXCR4^{R334X}$ . Both affected daughters (WHIM-10 and WHIM-11) tested positive, whereas the unaffected husband and third unaffected daughter tested negative. Surprisingly, two independent whole blood samples from



**Figure 1. Long-Term Clinical Remission of WHIM Syndrome and Evidence for Somatic Mosaicism in Patient WHIM-09**

(A) Family pedigree. Squares, males; circles, females; shaded symbols, WHIM syndrome; arrow, index patient WHIM-09.

(B) Spontaneous and complete remission of warts in patient WHIM-09. According to patient WHIM-09, through her fourth decade of life she had had extensive warts on her hands, similar to her daughters (illustrated here for 24-year-old daughter WHIM-11), that spontaneously resolved.

(C) Spontaneous sustained correction of neutropenia and monocytopenia in patient WHIM-09. WBC, white blood cell count; ANC, absolute neutrophil count; AMC, absolute monocyte count; ALC, absolute lymphocyte count. Arrows indicate age at splenectomy; horizontal lines indicate normal range for each cell type. Note, x axis is discontinuous to show pre/post splenectomy results more clearly.

(legend continued on next page)

WHIM-09, in which the leukocyte content was composed mostly of neutrophils and monocytes (Figures 1E and 1F), also tested negative for the mutation. In contrast, DNA from WHIM-09 PBMCs, which were composed mostly of lymphocytes, as well as from a lymphoblastoid cell line generated from WHIM-09 PBMCs both tested positive for CXCR4<sup>R334X</sup> (Figure 1E). The mutation was also not detectable by direct sequencing of whole blood cell DNA from WHIM-09, whereas whole blood cell samples from both daughters, WHIM-10 and WHIM-11, were both positive (Figure 1F). In contrast, DNA samples from a cheek swab and fibroblasts cultured from a skin biopsy from WHIM-09 were both heterozygous for CXCR4<sup>R334X</sup> (Figures 1E and 1G), defining her as a somatic genetic mosaic. A WHIM pedigree with germ-line/somatic genetic mosaicism, where the mother of two affected children was hematologically normal, has previously been described (Hernandez et al., 2003), but was unlikely to apply to WHIM-09 since she had been markedly symptomatic with severe neutropenia and myelokathexis as a child and cutaneous warts as an adult (Krill et al., 1964; Zuelzer, 1964). Therefore, we investigated whether a genetic reversion had occurred.

### Chromothripsis as the Mechanism for Genetic Mosaicism in Patient WHIM-09

The PCR-RFLP assay and DNA sequencing results may both be explained by either reversion or deletion in patient leukocytes of the mutant nucleotide that defines the CXCR4<sup>R334X</sup> allele. To address this, we first performed cytogenetic analysis of bone marrow cells. In all 20 metaphase cells examined, one copy of chromosome 2, where CXCR4 is located, was acrocentric rather than submetacentric and was significantly shorter than its normal homolog (Figure 2A). All other chromosomes appeared normal. In addition to deletions, the banding pattern of the abnormal chromosome 2 suggested the presence of inversions. Fluorescence in situ hybridization (FISH) revealed that the anaplastic lymphoma receptor tyrosine kinase (ALK) gene, at 2p23 on the normal chromosome 2, was on the long arm of the abnormal chromosome 2 (Figure 2B). FISH also demonstrated that the abnormal chromosome 2 had a portion of the centromere inverted into the long arm and the N-myc gene (MYCN), normally on 2p24, was absent (Figure 2C). Almost all of the interphase bone marrow cells with polymorphic nuclei, consistent with neutrophils, had the abnormal hybridization patterns of the abnormal chromosome 2 (Figure 2C), whereas cells with round nuclei had a significantly lower percentage of the abnormal chromosome (data not shown). An immunohistochemical stain of ALK activity on a bone marrow core biopsy section

did not reveal abnormal activation of the enzyme despite its abnormal chromosomal location (Figure S1).

To define the chromosomal abnormalities with greater resolution, microarray was performed using patient bone marrow cell and cultured fibroblast DNA. This revealed that the abnormal chromosome 2 was ~35 Mb shorter than usual due to seven large deletions (Figure S2A), as demonstrated by loss of heterozygosity in these regions and by an abrupt change in the relative copy number from 2 to 1. One of the deletions involved the MYCN gene, confirming the FISH results. In addition, one of the deletions included the position of the CXCR4 gene (Figure S2B). This had made the patient hemizygous for CXCR4 in cells having the abnormal chromosome 2 with loss of the CXCR4<sup>R334X</sup> mutation as well as loss of one copy of 163 other annotated genes (Table S1). Thus, development of hemizygosity of CXCR4, with only the wild-type copy remaining, explained why the CXCR4<sup>R334X</sup> mutation was not detectable in neutrophil DNA by either DNA sequencing or PCR-RFLP. However, this potentially fortuitous deletion was clearly only a small part of a much larger and more complex genetic event.

We next performed whole genome sequencing (WGS) for single nucleotide definition of the event, comparing blood neutrophil and skin fibroblast DNA from WHIM-09 to each other and to the standard human genome sequence in GenBank (version hg19). We obtained an average of 40× coverage of both samples (database of Genotypes and Phenotypes [dbGaP] accession number phs000856.v1.p1. [[http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study\\_id=phs000856.v1.p1](http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000856.v1.p1)]) and an analysis was performed to locate the DNA reads that had homology with multiple distinct areas of chromosome 2. This technique allowed precise base pair level identification of the inversions and deletions that had occurred on one copy of chromosome 2 (Table S2). No other large deletions or inversions were detected in the genome.

The changes seen in the microarray were confirmed by WGS but were actually more complex than had been initially suspected. The derivative chromosome was composed of 18 remaining pieces arranged in random orientation and in a random order, characteristic of chromothripsis (Figure 3A). A circular plot of the connections revealed by paired end WGS of the neutrophil DNA is shown in Figure 3B. The abnormal derivative chromosome, modeled using the breakpoints and connections (Figure 3C), revealed a predicted structural banding pattern that was identical to what was observed by cytogenetic analysis. We also developed primer pairs spanning four

(D) Normalization of bone marrow pathology in patient WHIM-09. A representative high magnification (500×) Wright-Giemsa stain of the bone marrow aspirate is shown for the index patient WHIM-09 in 1963 (Zuelzer, 1964) (reproduced with permission) and in 2013 at ages 9 and 59, respectively. Arrows in left image, eyeglass nuclei in neutrophils.

(E) PCR-BstUI restriction fragment length polymorphism (BstUI) analysis of genomic DNA. HD, healthy donor; WHIM-09, index patient; Cheek, cheek swab cells; PBMC, peripheral blood mononuclear cells; Fb, fibroblast; LCL, lymphoblastoid cell line; WB, whole blood; WT, wild-type allele; WHIM, CXCR4<sup>R334X</sup> allele that causes WHIM syndrome.

(F) Sanger DNA sequencing analysis of whole blood DNA for affected family members in the region near nucleotide position 1,000 (vertical line), the site of WHIM mutation CXCR4<sup>R334X</sup> (1,000 C → T). Blue, C; Red, T; Fb, fibroblast; PMN, polymorphonuclear leukocyte (PMN).

(G) Sanger DNA sequencing analysis of DNA from purified peripheral blood neutrophils, cells obtained from a buccal swab (cheek) and cultured skin fibroblasts for patient WHIM-09 in the same region as (F).

**Table 1. Distribution of Leukocyte Subsets in the Peripheral Blood of Index Patient WHIM-09 in Clinical Remission from WHIM Syndrome and Her Two Affected Daughters, WHIM-10 and WHIM-11**

Leukocyte Subset	WHIM-09	WHIM-10	WHIM-11	Normal Adult Reference Range <sup>c</sup>
Neutrophil	9,980 <sup>a</sup>	380 <sup>b</sup>	530 <sup>b</sup>	1,560–6,130
CD3 <sup>+</sup> T cell	1,736	382 <sup>b</sup>	218 <sup>b</sup>	714–2,266
CD4 <sup>+</sup> T cell	1,389	317 <sup>b</sup>	155 <sup>b</sup>	359–1,565
CD4 <sup>+</sup> CD45RA <sup>+</sup> naive T cell	211 <sup>b</sup>	44 <sup>b</sup>	12 <sup>b</sup>	454–733
CD4 <sup>+</sup> CD45RO <sup>+</sup> memory T cell	1,178 <sup>a</sup>	273	143 <sup>b</sup>	219–1,048
CD8 <sup>+</sup> T cell	222	41 <sup>b</sup>	45 <sup>b</sup>	178–853
CD8 <sup>+</sup> CD45RA <sup>+</sup> naive T cell	44 <sup>b</sup>	22 <sup>b</sup>	14 <sup>b</sup>	231–371
CD8 <sup>+</sup> CD45RO <sup>+</sup> memory T cell	178 <sup>a</sup>	19 <sup>b</sup>	31 <sup>b</sup>	57–130
CD3 <sup>-</sup> CD56 <sup>+</sup> NK cell	1,270 <sup>a</sup>	32 <sup>b</sup>	51 <sup>b</sup>	126–729
CD19 <sup>+</sup> B cell	18 <sup>b</sup>	5 <sup>b</sup>	10 <sup>b</sup>	61–329
CD19 <sup>+</sup> CD27 <sup>+</sup> memory B cell	2 <sup>b</sup>	1 <sup>b</sup>	1 <sup>b</sup>	12–68
CD19 <sup>+</sup> CD27 <sup>-</sup> B cell	16 <sup>b</sup>	4 <sup>b</sup>	9 <sup>b</sup>	90–176
CD19 <sup>+</sup> CD27 <sup>-</sup> IgD <sup>+</sup> IgM <sup>+</sup> transitional/naive B cell	12 <sup>b</sup>	3 <sup>b</sup>	4 <sup>b</sup>	42–85
CD19 <sup>+</sup> CD27 <sup>-</sup> IgD <sup>-</sup> IgM <sup>+</sup> immature B cell	4	1 <sup>b</sup>	1 <sup>b</sup>	2–10
CD14 <sup>+</sup> CD16 <sup>-</sup> classical monocyte	1,344 <sup>a</sup>	35 <sup>b</sup>	117 <sup>b</sup>	371–539
CD14 <sup>+</sup> CD16 <sup>+</sup> inflammatory monocyte	235 <sup>a</sup>	5 <sup>b</sup>	23	14–30

Data are presented as absolute numbers of cells having the indicated immunophenotype per microliter of whole blood.

<sup>a</sup>Value above the upper limit of normal.

<sup>b</sup>Value below the lower limit of normal.

<sup>c</sup>Based on the values of 11–40 healthy blood donors seen at the NIH Clinical Center.

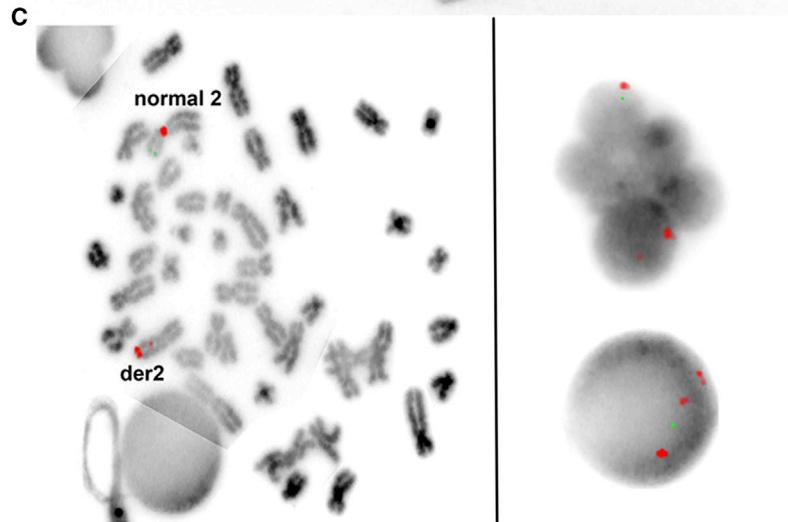
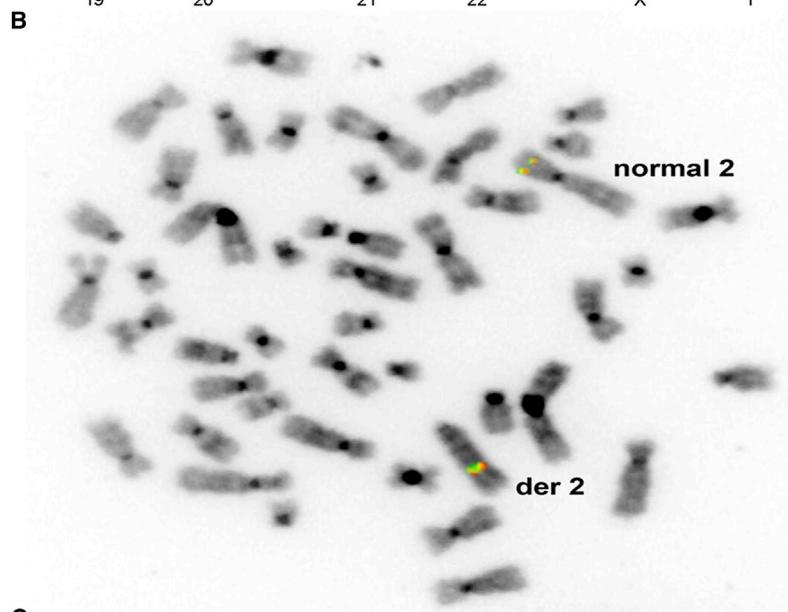
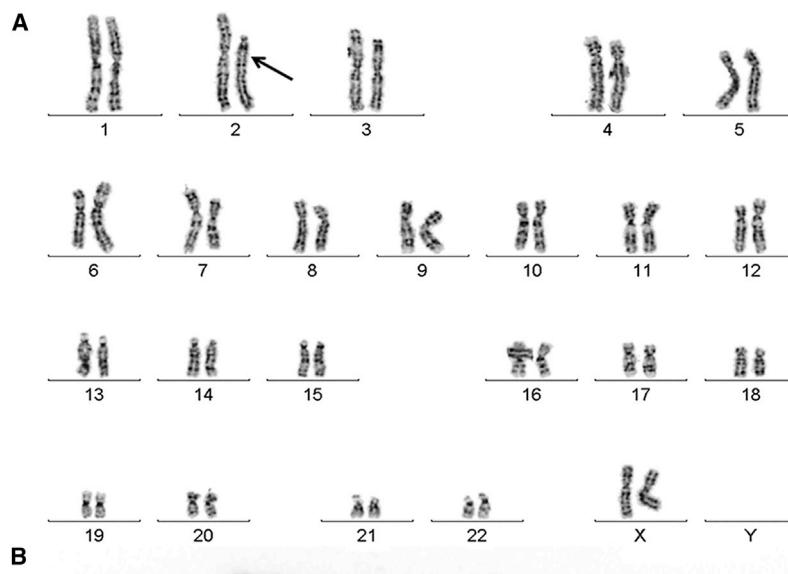
of the unique chromothriptic boundaries and demonstrated experimentally that they generated the predicted rearranged product from WHIM-09 neutrophil DNA but not from healthy donor DNA (data not shown). By analyzing the rearrangement breakpoints for the creation of novel fusion genes, we found two possibilities, fusions of *MFSD2B* with *LOC285000* and *POTEKP* with *NCKAP5* (Figure S3); however, these genes are unlikely to be expressed or functional because each lacks a promoter and transcription initiation site, and major regions of the potentially fused genes are deleted. Thus, the most likely explanation for the patient's clinical improvement was haploinsufficiency for one or more of the 164 genes affected by chromothripsy.

### Myeloid/Lymphoid Mosaicism Established by Chromothripsy in an HSC from Patient WHIM-09

Chromothripsy-specific primers did not generate a detectable product by PCR when DNA prepared from the archival spleen sample from patient WHIM-09 was amplified, suggesting that the chromothriptic event occurred after age 9, when splenectomy was performed, but before age 35, when the rise in WBC was already underway (Figure 4A). The emergence of uniformly chromothriptic neutrophils, as suggested by WGS, PCR, and bone marrow cytogenetics and FISH, implied that at least some patient HSCs must also be chromothriptic. We verified this with CD34<sup>+</sup> leukocytes cultured from peripheral blood, as well as with CD38<sup>-</sup>CD90<sup>+</sup>CD45RA<sup>-</sup> HSCs FACS-sorted from patient bone marrow (Figures 4B, 4C, and S4) that both appeared to be markedly deficient in or lack the CXCR4<sup>R334X</sup> allele by the *Bst*UI PCR-RFLP assay. Consistent with this, CD38<sup>+</sup>CD135<sup>+</sup>CD45RA<sup>-</sup> common myeloid precursor cells (CMP) and CD38<sup>+</sup>CD135<sup>+</sup>CD45RA<sup>+</sup> granulocyte-monocyte precursor cells (GMP) sorted from patient bone marrow also lacked the CXCR4<sup>R334X</sup> allele by the PCR-RFLP assay, whereas FACS-sorted CD45<sup>+</sup> CD34<sup>+</sup> CD38<sup>+</sup> CD10<sup>+</sup> common lymphoid precursors (CLP) remained CXCR4<sup>+/R334X</sup> (Figure 4D, lower panel). This pattern was confirmed by PCR using chromothripsy-specific primers and DNA from HSCs, CMPs, and CLPs (Figure 4D, upper panel). Consistent with this, we found that all mature myeloid cell types tested in both blood and bone marrow lacked or were markedly deficient in the CXCR4<sup>R334X</sup> WHIM allele, whereas all mature lymphoid cell types tested in peripheral blood and bone marrow were heterozygous CXCR4<sup>+/R334X</sup> (Figures 4B and 4C). This aligns with the cytogenetic results presented in Figure 2C. Further, we found 100% of EBV-transformed B cell lines ( $n = 10$ ) prepared from patient blood lacked the chromothriptic chromosome (data not shown). As expected, erythroid precursors, which derive from CMPs, also lacked the CXCR4<sup>R334X</sup> allele as demonstrated by *Bst*UI PCR-RFLP analysis of Burst Forming Unit-Erythroid (BFU-E) colonies generated from patient CD34<sup>+</sup> cells cultured from PBMCs ex vivo (Figure 4E). Thus, the combined evidence suggests that an HSC underwent chromothripsy and selectively repopulated the myeloid lineage (including the erythroid lineage), but not the lymphoid lineage (Figure 4F).

### Cxcr4 Haploinsufficiency Enhances Engraftment during Bone Marrow Transplantation

To test whether deletion of the disease allele CXCR4<sup>R334X</sup> alone, and not any of the other 163 genes deleted by chromothripsy would be sufficient to confer a selective advantage to an HSC in the context of WHIM syndrome, we performed competitive repopulation experiments using a mouse model of WHIM syndrome (*Cxcr4*<sup>+/S338X</sup>), hemizygous *Cxcr4* mice (*Cxcr4*<sup>+/-</sup>), the same as the HSC and myeloid cell CXCR4 genotype of patient WHIM-09), and wild-type (*Cxcr4*<sup>++</sup>) mice as donors and recipients. Leukocytes from these strains were specifically marked with either CD45.1 or CD45.2 or both, allowing tracking of transplanted donor cell fate in vivo by FACS. When  $\sim 2.5 \times 10^6$  total donor bone marrow cells from both *Cxcr4*<sup>+/S338X</sup> and *Cxcr4*<sup>+/-</sup> mice were mixed together in

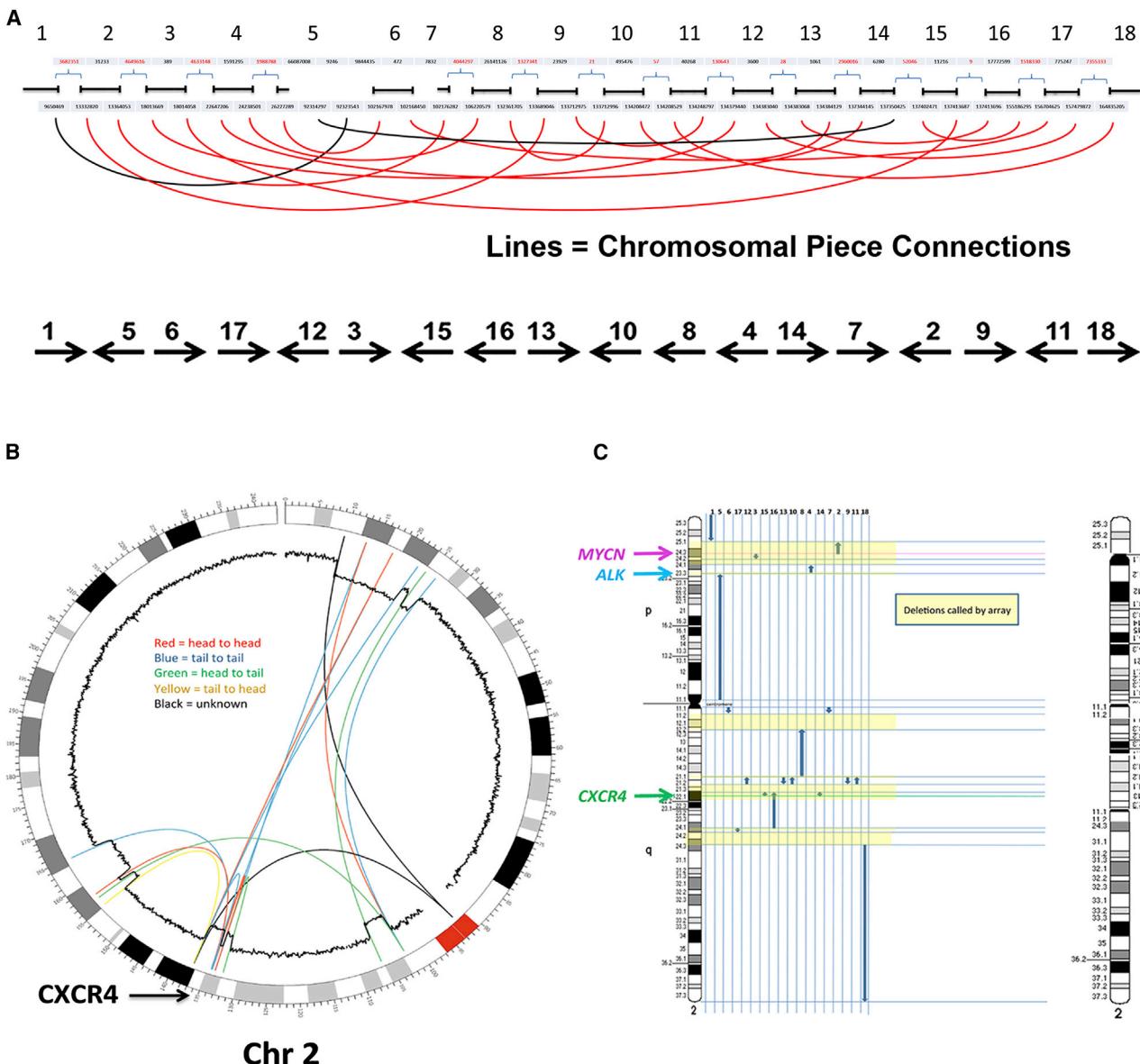


**Figure 2. Massive Deletion and Rearrangement of One Copy of Chromosome 2, the Location of CXCR4, in Patient WHIM-09**

Cytogenetics and fluorescence in situ hybridization (FISH) analyses were performed on cultured bone marrow cells. (A) Cytogenetics. Karyogram showing a short acrocentric chromosome 2 (arrow) with abnormal banding pattern suggesting deletions and inversions observed in all 20 metaphase cells analyzed from bone marrow.

(B) FISH with intact ALK break apart probe (Abbott Molecular) signal at 2p23 on normal chromosome 2 and on long arm of abnormal chromosome 2 (der 2).

(C) FISH with *MYCN* and CEP 2 probe set (Abbott Molecular). Left: metaphase cell showing the normal chromosome 2 with intact centromere signal (red) and *MYCN* signal (green) at 2p24. The abnormal chromosome 2 has a portion of its centromere inverted into the long arm splitting the red signal; the green *MYCN* signal is absent. Right: polymorphonuclear and round interphase nuclei showing the abnormal hybridization pattern with one green *MYCN* signal and three red CEP 2 signals. One of the red signals is smaller than the other two and is close to one of them. See also Figure S1.

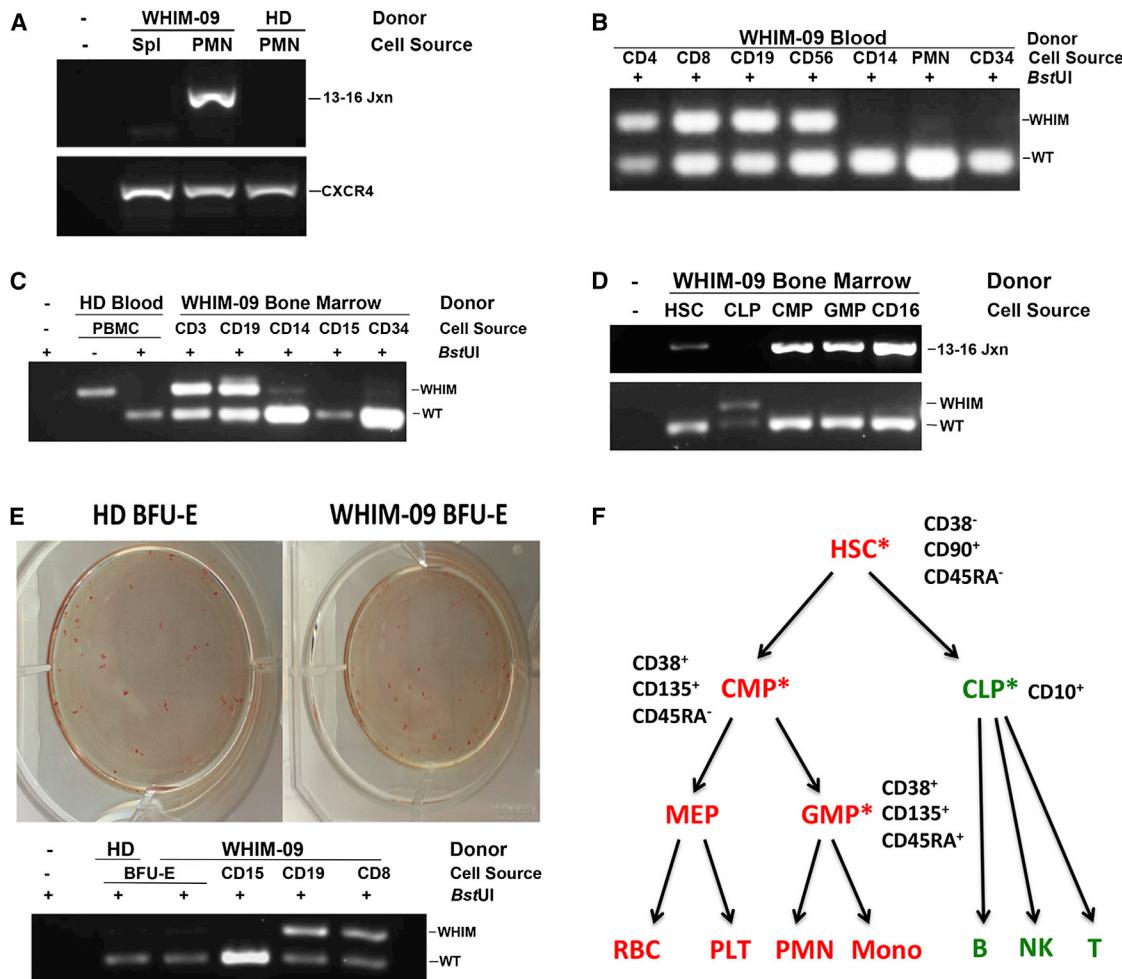


**Figure 3. Chromothripsis as the Mechanism for Loss of CXCR4<sup>R334X</sup> in Patient WHIM-09**

Purified neutrophil and cultured skin fibroblast DNA from patient WHIM-09 was isolated and subjected to whole genome sequencing with paired-end analysis. (A) Linear non-proportional plot of the abnormal copy of chromosome 2 in patient WHIM-09 labeled from the p arm telomere (0) to the q arm telomere (243) in megabases with the 18 remaining pieces arranged in their numeric order (top). Connections between these pieces are depicted by the curved lines. Note that some connections were poorly defined because of the involvement of repetitive centromeric sequence (black lines). The order and orientation of the 18 remaining pieces in the derivative chromosome are indicated at the bottom.

(B) Circos plot of chromosome 2 and its normal Giemsa cytogenetic banding pattern labeled from the p arm telomere (0) to the q arm telomere (~240) in megabases. Large pieces of chromosome 2 were missing from patient WHIM-09 neutrophil DNA and the 18 remaining pieces were arranged in random order. Connections between these pieces and their orientation (inset) are depicted by the colored lines. See Figure S2 and Tables S1 and S2 for additional details. Note that two connections were poorly defined because of the involvement of repetitive centromeric sequence. The inner circular trace is the copy number variation data derived from microarray analysis (Figures S2A and S2B). Note that the sites of connections derived from the paired-end sequencing analysis closely match the sites where copy number variation abruptly falls from 2 to 1. The location of CXCR4 is indicated at lower left.

(C) Derivative chromothriptic chromosome 2. Ideogram of intact chromosome 2 (left) and a model of the Giemsa cytogenetic banding pattern and the order and orientation of the 18 pieces with the deletions called by microarray shaded in yellow (right) are shown. The resultant remaining chromosome 2 banding pattern predicted by whole genome sequencing closely matches that seen by cytogenetic analysis. Note the location of CXCR4 at 2q22.1 in one of the deleted segments. See also Figure S3.



**Figure 4. Chromothriptic CXCR4-Haploinsufficient HSC Replacement of the Myeloid Lineage, but Not the Lymphoid Lineage, Is Associated with Clinical Remission in Patient WHIM-09**

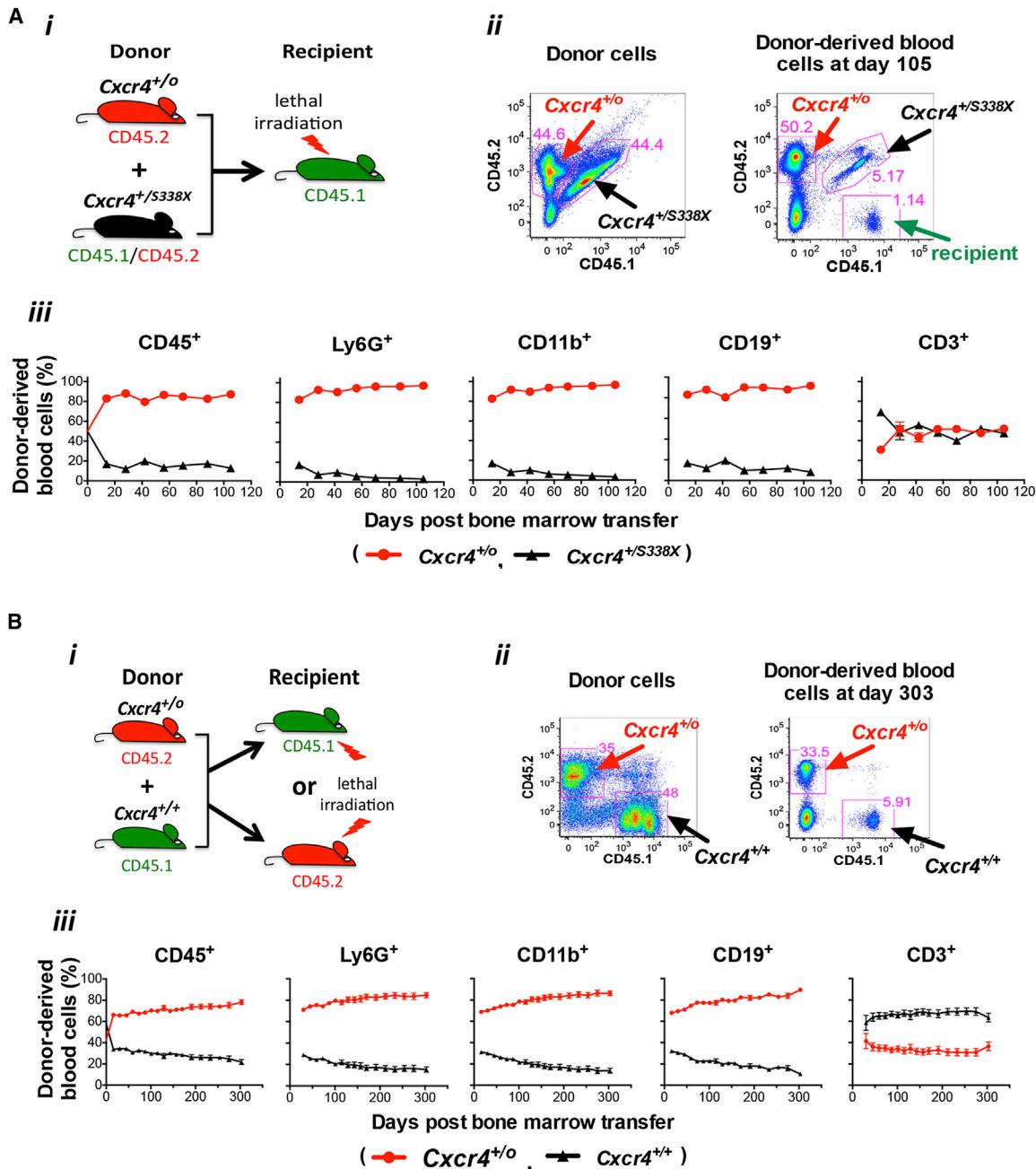
(A–E) Representative results from a *Bst*UI PCR-restriction fragment length polymorphism assay (*Bst*UI), designed to distinguish the wild-type CXCR4 allele (WT) from the CXCR4<sup>R334X</sup> WHIM allele (WHIM), as well as from a PCR assay specific for the chromothriptic chromosome (13–16 Jxn). PCR was performed on DNA obtained from the indicated donor leukocyte subsets purified either from blood using magnetic bead purification (B and E) or from a bone marrow aspirate using flow cytometric sorting (C and D). DNA was also prepared from archived WHIM-09 spleen and compared with peripheral blood PMN DNA (A), as well as from Burst-forming Unit-Erythroid colonies and compared with blood leukocyte subsets (E). WHIM-09, index patient; HD, healthy donor; Spl, spleen; PMN, polymorphonuclear leukocytes; 13–16 Jxn, PCR product specific for the chromothriptic junction between segments 13 and 16 of the chromothriptic chromosome of patient WHIM-09; PBMC, peripheral blood mononuclear cells; CD4, purified CD4<sup>+</sup> T cells; CD8, purified CD8<sup>+</sup> T cells; CD19, purified CD19<sup>+</sup> B cells; CD56, purified CD56<sup>+</sup> natural killer cells; CD14, purified CD14<sup>+</sup> monocytes; CD34, purified CD34<sup>+</sup> hematopoietic cells; CD3, purified CD3<sup>+</sup> T cells; CD15, purified CD15<sup>+</sup> neutrophils; CD16, purified CD16<sup>+</sup> neutrophils; HSC, hematopoietic stem cells; CLP, common lymphoid precursor; CMP, common myeloid precursor; GMP, granulocyte/monocyte precursor MEP, megakaryocyte-erythroid precursor; BFU-E, Burst-forming Unit-Erythroid; CXCR4, CXCR4 amplicon not digested with *Bst*UI.

(F) Summary of myeloid/lymphoid mosaicism for CXCR4<sup>R334X</sup> in patient WHIM-09. The immunophenotype used to purify each cell type from enriched CD34<sup>+</sup>CD45<sup>+</sup> cells is summarized next to each cell type shown. Red, negative for CXCR4<sup>R334X</sup>; green, positive for CXCR4<sup>R334X</sup>; asterisks, purified cell types directly analyzed by PCR-RFLP for the WHIM mutation.

See also Figure S4.

equal proportions (~50%:50%) and transplanted into lethally irradiated *Cxcr4*<sup>+/+</sup> recipients, by day 105 post-transplantation the percentage of *Cxcr4*<sup>+/S338X</sup> neutrophils, monocytes, and B cells detectable in the peripheral blood had declined to ~5%, whereas the percentage of each of the corresponding *Cxcr4*<sup>+/0</sup> leukocyte subsets had increased to ~95% (Figure 5A). The rate at which the proportion of donor-derived blood cells

diverged in frequency from the input ratio of 50:50 occurred in two phases, a rapid early phase within 2 weeks and a slow late phase evident by 2 weeks after transplantation. The inflection point could be earlier than 2 weeks, since this is the first time point when data were collected to allow for recovery after transplantation. Importantly, there was no difference in the baseline HSC frequency in bone marrow for *Cxcr4*<sup>+/0</sup>



**Figure 5. *Cxcr4* haploinsufficiency enhances HSC engraftment during mouse bone marrow transplantation**

Two types of competitive bone marrow transplantation experiments were performed: (A) *Cxcr4<sup>+/o</sup>* versus *Cxcr4<sup>+/S338X</sup>* (mouse model of WHIM syndrome), and (B) *Cxcr4<sup>+/o</sup>* versus *Cxcr4<sup>+/+</sup>*. (i) Experimental design. (ii) Representative flow cytometry plots demonstrating the relative contributions of CD45 congenic markers in mixed donor bone marrow prior to transplantation (left panel) and in blood after bone marrow transplantation (right panel) in a single mouse. (iii) Cell frequency data for the leukocyte subsets indicated at the top of each panel, presented as the mean ± SEM percentage (%) of total donor-derived cells for each subset ( $n = 10$  mice per data point). SEM was <5% of the mean in all cases and therefore is not visible for most data points. Results were verified in one and two additional independent experiments for (A) and (B), respectively.

See also Figure S5.

and *Cxcr4<sup>+/S338X</sup>* donor mice, so the number of HSCs transplanted was similar (Figure S5). Since the ANC in the peripheral blood of *Cxcr4<sup>+/S338X</sup>* donor mice is only ~50% less than the corresponding value in both *Cxcr4<sup>+/+</sup>* mice (Bala-

banian et al., 2012; J.-L.G., M.S., E.Y., and P.M.M., data not shown) and *Cxcr4<sup>+/o</sup>* mice (J.-L.G., unpublished data), the extreme skewing of these transplantation results in the blood suggested that hemizygous *Cxcr4* (*Cxcr4<sup>+/o</sup>*) HSCs may have

a selective advantage over WHIM ( $\text{Cxcr}4^{+/S338X}$ ) HSCs for engraftment in this system. Consistent with this, in the bone marrow, as in the blood, we observed a greater content of mature  $\text{Cxcr}4^{+/0}$  leukocytes compared to mature  $\text{Cxcr}4^{+/S338X}$  leukocytes (data not shown). No difference was observed for the frequencies of  $\text{Cxcr}4^{+/0}$  and  $\text{Cxcr}4^{+/S338X}$  T cells in the blood, presumably owing to homeostatic proliferation of mature T cells present in donor bone marrow transferred into irradiated hosts.

To test whether there might be a general selective advantage of  $\text{Cxcr}4$  hemizygosity in transplantation, we performed competitive repopulation experiments by mixing  $\sim 2.5 \times 10^6$  total donor bone marrow cells each from  $\text{Cxcr}4^{+/+}$  and  $\text{Cxcr}4^{+/0}$  mice and transplanting the mixture into lethally irradiated  $\text{Cxcr}4^{+/+}$  recipients. The baseline HSC frequency in bone marrow was the same for  $\text{Cxcr}4^{+/0}$  and  $\text{Cxcr}4^{+/+}$  donor mice (Figure S5). In this case, the input ratio was skewed slightly in favor of the  $\text{Cxcr}4^{+/+}$  cells over the  $\text{Cxcr}4^{+/0}$  cells (58%:42%). Nevertheless, by day 303 post-transplantation, when the animals were sacrificed, the percentage of  $\text{Cxcr}4^{+/+}$  neutrophils, monocytes and B cells detectable in the peripheral blood had declined to less than  $\sim 15\%$ , whereas the percentage of each of the corresponding  $\text{Cxcr}4^{+/0}$  leukocyte subsets had increased to  $\sim 85\%$  with similar kinetics as for the competition with  $\text{Cxcr}4^{+/S338X}$  bone marrow (Figure 5B). The same effect was observed whether the irradiated recipient mouse was CD45.1, CD45.2, or  $\text{Cxcr}4^{+/0}$  (data not shown) and whether the donor bone marrows were depleted of lineage-positive cells (Figure S6).

To investigate the mechanism for the apparent competitive advantage of  $\text{Cxcr}4^{+/0}$  over  $\text{Cxcr}4^{+/+}$  bone marrow cells for reconstituting the blood, we first measured the proliferative status of the corresponding HSCs in vivo by BrdU incorporation early after transplantation (day 7) (Figure 6A). The results indicated  $\sim 20\%$  greater frequency of  $\text{BrdU}^+$   $\text{Cxcr}4^{+/0}$  HSCs as compared to  $\text{Cxcr}4^{+/+}$  HSCs suggesting a proliferative advantage of  $\text{Cxcr}4^{+/0}$  HSCs in the bone marrow. To test whether differential retention of leukocytes in the bone marrow might also contribute to the skewed distribution of mature leukocytes in the blood, we sacrificed mice at day 303 after competitive transplantation with  $\text{Cxcr}4^{+/0}$  and  $\text{Cxcr}4^{+/+}$  bone marrow cells (the same combination of donor bone marrow cells as were analyzed in the proliferation experiments). The results showed that  $\text{Cxcr}4^{+/0}$  HSCs, HPCs (hematopoietic progenitor cells), and total CD45<sup>+</sup> cells (mostly mature leukocytes) predominated in the bone marrow by the same 4:1 ratio over the corresponding  $\text{Cxcr}4^{+/+}$  cells as was found for mature leukocytes in the blood. Thus, the predominance of mature  $\text{Cxcr}4^{+/0}$  over  $\text{Cxcr}4^{+/+}$  leukocytes in the blood of competitively transplanted mice cannot be simply explained by low retention of mature  $\text{Cxcr}4^{+/0}$  leukocytes relative to the retention of mature  $\text{Cxcr}4^{+/+}$  leukocytes in the bone marrow. Moreover, the results at this very late time point clarify that enhanced  $\text{Cxcr}4^{+/0}$  HSC proliferation does not result in long-term depletion of HSCs. The frequency distribution of stem and progenitor subtypes was the same for each genotype, indicating no block to progenitor cell differentiation (Figure 6B).

Finally, when  $\text{Cxcr}4^{+/0}$  bone marrow was used in competitive repopulation assays with either  $\text{Cxcr}4^{+/+}$  or  $\text{Cxcr}4^{+/S338X}$ ,

we found in both cases that the total white blood cell and differential counts in the blood post-transplantation were not significantly altered as compared to reference values for C57Bl/6J mice (<http://phenome.jax.org>), suggesting that  $\text{Cxcr}4^{+/0}$  genotype does not predispose to myelodysplasia or leukemia in the setting of this transplantation model (Figure S7). Moreover, there is no evidence that  $\text{Cxcr}4^{+/0}$  mice have a constitutive predisposition to myelodysplasia (data not shown).

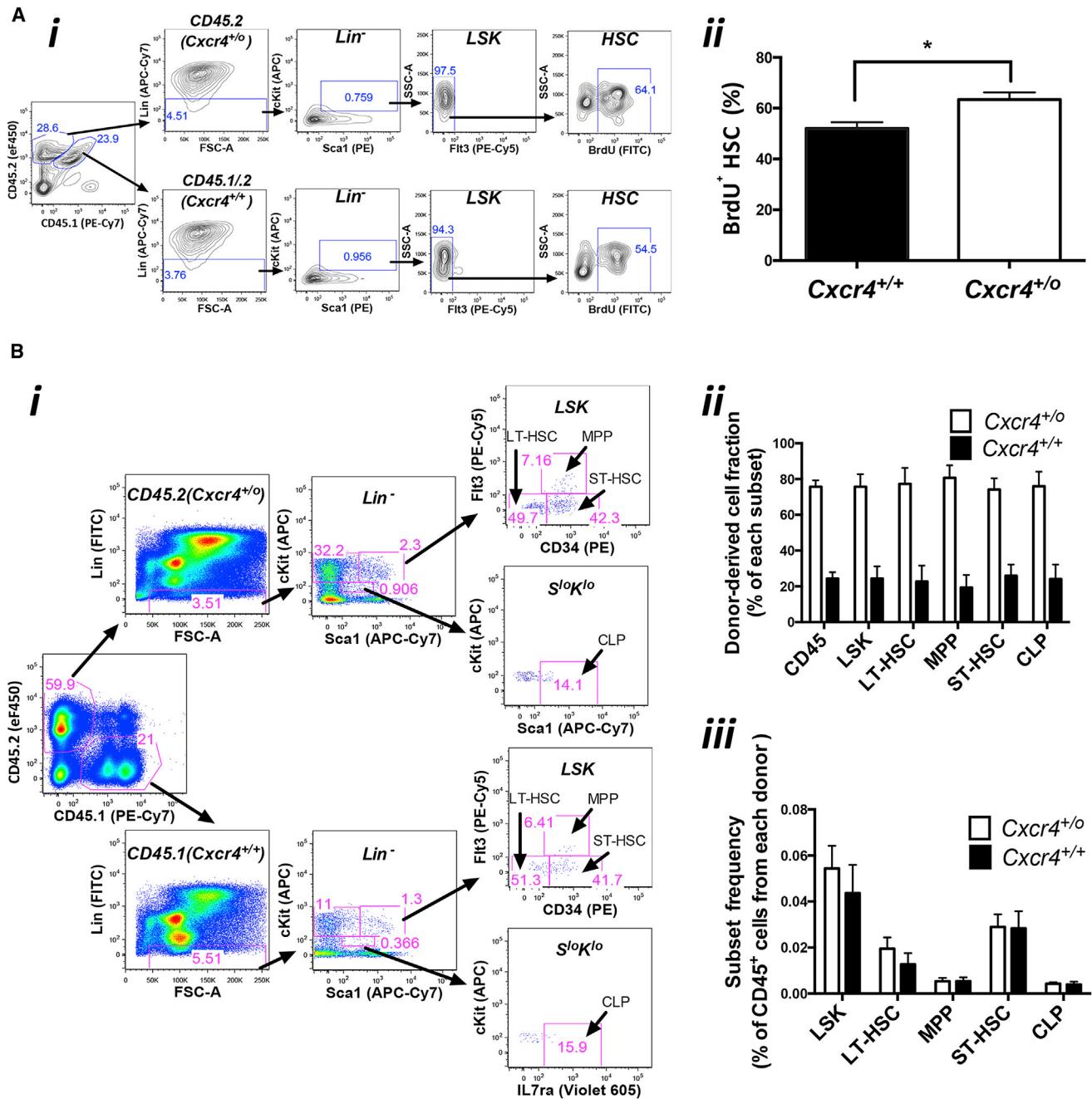
## DISCUSSION

Our analysis of this unprecedented experiment of nature in a single patient provides the first evidence that chromothripsis, a newly described form of genetic instability, may result in clinical benefit, in particular, cure of an inherited disease. This patient also provides the second example of chromothripsis occurring in a stem cell still capable of terminal differentiation, (Kloosterman et al., 2011) as well as one of the few examples of spontaneous genetic cure of a Mendelian condition (Hirschhorn, 2003).

We use the word “cure” to refer to the patient’s clinical outcome because she had met all four of the acronymic clinical criteria diagnostic for WHIM syndrome through her fourth decade of life, but has fulfilled none except for mild hypogammaglobulinemia since then ( $\sim 20$  years, to date). Moreover, the disease allele  $\text{CXCR}4^{R334X}$  is no longer detectable in key cell types that drive the disease: HSCs, neutrophils, and monocytes. To our knowledge, patient WHIM-09’s spontaneous long-term complete remission of warts without treatment is unprecedented in WHIM syndrome (Al Ustwani et al., 2014; Beausant Cohen et al., 2012; Kawai and Malech, 2009; Tassone et al., 2009; Wetzler et al., 1990). Since she remains lymphopenic (B and naive T lymphocytes), this suggests that the myeloid arm of the immune system, probably through monocytes or monocyte-derived cells, plays an essential role in HPV clearance.

The combined evidence suggests that an HSC in WHIM-09 underwent chromothripsis between the second and fourth decade of life and selectively repopulated the myeloid lineage, but not the lymphoid lineage. This pattern of hematologic mosaicism implies that the chromothriptic changes precluded differentiation of HSCs to CLPs but not to CMPs. Thus, the mechanism for maintenance of the B cell lineage, which is WHIM in the patient, is unclear. Possibilities include differentiation from a small population of persistent WHIM HSCs below the level of detection of our assays, or by differentiation of self-sustaining or very long-lived WHIM CLPs. In contrast, the T cell lineage may simply have been maintained by homeostatic proliferation of the pre-chromothriptic WHIM T cell repertoire.

Chromothripsis is a complex chromosomal catastrophe that is thought to occur all at once in one cell (Stephens et al., 2011). If the affected cell dies, chromothripsis is clinically silent and undetectable, therefore the true frequency with which it occurs cannot be determined. If the chromothriptic cell acquires a strong selective advantage, it may emerge as a readily detectable, clinically apparent clonal population harboring a single



**Figure 6. Cxcr4 Haploinsufficiency Enhances HSC Proliferation and Engraftment in a Mouse Bone Marrow Transplantation Model**

(A) Proliferation. Donor bone marrow cells with a  $Cxcr4^{+/o}$  genotype on a homozygous CD45.2 background were mixed with donor bone marrow cells with a  $Cxcr4^{+/+}$  genotype on a heterozygous CD45.1/CD45.2 background (47:53) and injected intravenously into a lethally irradiated CD45.1 recipient mouse. Six days after bone marrow transplantation, each mouse was given 1.25 mg of BrdU intraperitoneally (i.p.). Twenty hours later, the mice were euthanized for HSC proliferation analysis. (i) Gating scheme for BrdU<sup>+</sup> HSCs. Bone marrow cells were first gated with CD45.2 ( $Cxcr4^{+/o}$ ) and CD45.1/CD45.2 ( $Cxcr4^{+/+}$ ), then HSCs were gated as Flt3<sup>-</sup> Lin<sup>-</sup> Sca1<sup>+</sup> c-Kit<sup>+</sup> (Flt3<sup>-</sup> LSK), which includes long-term and short-term HSCs, and BrdU<sup>+</sup> cells were quantitated. (ii) Percentage of BrdU<sup>+</sup> HSCs in each donor. Data are expressed as mean  $\pm$  SEM from four mice. The experiment was repeated once with similar results.

(B) Long-term engraftment and differentiation. Bone marrow cells from donor mice with a  $Cxcr4^{+/o}$  genotype on a CD45.2 background were mixed with bone marrow cells from donors with a  $Cxcr4^{+/+}$  genotype on a CD45.1 background (42:58) and the mixed cells were injected intravenously into lethally irradiated recipient mice. Bone marrow was harvested 303 days later. (i) Gating scheme for long-term HSC (LT-HSC: CD34<sup>+</sup> Flt3<sup>-</sup> Lin<sup>-</sup> Sca1<sup>+</sup> c-Kit<sup>+</sup>), short-term HSC (ST-HSC: CD34<sup>+</sup> Flt3<sup>-</sup> Lin<sup>-</sup> Sca1<sup>+</sup> c-Kit<sup>+</sup>), multipotent progenitors (MPP: CD34<sup>+</sup> Flt3<sup>-</sup> Lin<sup>-</sup> Sca1<sup>+</sup> c-Kit<sup>+</sup>), and common lymphoid progenitors (CLP: IL7ra<sup>+</sup> Lin<sup>-</sup> Sca1<sup>low</sup> c-Kit<sup>low</sup>). (ii) Long-term engraftment. (iii) Differentiation. The distribution frequency of bone marrow cell subsets is similar for  $Cxcr4^{+/o}$  compared to  $Cxcr4^{+/+}$  donor-derived cell populations. Each data point represents ten mice presented as the mean  $\pm$  SEM.

See also Figures S6 and S7.

pattern of deletions and rearrangements, resulting for example either in cancer, as previously reported (Jones and Jallepalli, 2012), or if the location of the event is fortuitous, in cure of a genetic condition, as occurred in patient WHIM-09. Although it is possible that deletion of the disease allele CXCR4<sup>R334X</sup> in the original chromothriptic HSC of patient WHIM-09, which rendered the cell CXCR4 haploinsufficient (CXCR4<sup>+/-</sup>), may have been sufficient to repopulate the myeloid lineage—as suggested by the strong HSC engraftment advantage of *Cxcr4* haploinsufficient mouse bone marrow cells in our competitive bone marrow transplantation experiments—other factors, particularly haploinsufficiency for one or more of the 163 other genes that were deleted by chromothripsy, may also have contributed. In this regard, at least three of the other 163 deleted genes, *DNMT3A* (Challen et al., 2012), *MYCN* (Laurenti et al., 2008), and *IL1R* (Orelio et al., 2009), have been reported to regulate hematopoiesis (Table S1).

That *Cxcr4* haploinsufficiency should provide a competitive engraftment advantage to HSCs is somewhat counterintuitive since *Cxcr4* has been well-documented to promote HSC homing to and retention in bone marrow (Broxmeyer et al., 2003a, 2003b, 2005; Dar et al., 2006; Kawai et al., 2007; Nie et al., 2008; Sugiyama et al., 2006). However, it has also been reported that CXCL12 secreted by CXCL12-abundant reticular (CAR) cells that contact HSCs in the bone marrow niche is able to enhance HSC quiescence by stimulating CXCR4 signaling (Nie et al., 2008; Sugiyama et al., 2006). The overall effect on engraftment of reducing signaling via CXCR4 haploinsufficiency has not been previously tested *in vivo* and cannot be confidently predicted given the diverse and contrasting roles of the receptor on HSCs as well as on mature leukocytes. Our mouse data confirmed *in vivo* that *Cxcr4* haploinsufficiency enhanced HSC proliferation while maintaining long-term hematopoiesis, therefore a competitive advantage acquired from enhanced proliferation and potentially other mechanisms by these cells may supercede any potential disadvantage they may have at the level of bone marrow retention and homing.

Additional work using more stringent engraftment protocols in mice and ultimately gene editing/transplantation trials in humans will be needed to precisely determine the impact of CXCR4 haploinsufficiency on HSC engraftment as well as on each of the parameters that regulate the physiologic steady-state levels of mature leukocytes and hematopoietic stem and progenitor cells in blood, bone marrow, and other hematopoietic compartments. Such studies may further elucidate precisely how patient WHIM-09 was cured, which may point to general applications in transplantation and gene therapy.

## EXPERIMENTAL PROCEDURES

### Patients

Consistent with the Declaration of Helsinki, all human subjects signed informed consent to participate in NIAID Institutional Review Board-approved clinical protocols. All subjects were studied at the NIH Clinical Center.

### Human Leukocyte Analysis

Leukocyte subsets were isolated from blood and bone marrow. PBMCs and bone marrow cells were cultured ex vivo to expand CD34<sup>+</sup> hematopoietic

progenitor cells and to generate BFU-E colonies. PBMCs were used to create lymphoblastoid cell lines (LCL) from Epstein-Barr virus (EBV)-transformed B cells.

### Genetic Analysis

*Bst*UI PCR-RFLP analysis was performed as previously described (Hernandez et al., 2003) using patient genomic DNA. The genome was also analyzed by PCR using primers spanning chromothriptic junctions of the derivative chromosome 2 of patient WHIM-09; by Sanger method sequencing; by standard cytogenetics and fluorescence *in situ* hybridization (FISH) of metaphase chromosomes (Quest Diagnostics) using Abbott Molecular probes (Abbott Molecular); by microarray analysis using the Affymetrix Cytoscan HD array (Affymetrix); and by whole genome sequencing (WGS). For WGS, paired-end patient WHIM-09 fibroblast and neutrophil DNA libraries were prepared using the TruSeq DNA protocol (Illumina) and were sequenced using a HiSeq 2000 next generation sequencer (Illumina) generating ~380 million reads per library of 2 × 100 bp paired-ends. Reads were mapped to build hg19 of the human genome to identify fusion junctions present in the neutrophil but not in the fibroblast samples.

### Transplantation Experiments

*Cxcr4* floxed mice (Strain 008767, B6.129P2-*Cxcr4*<sup>tm2Yzo</sup>/J) and Ella promoter driven Cre recombinase transgenic mice (Strain 003724, B6.FVB-Tg(Ella-cre) C5379Lmgd/J) were obtained from the Jackson Laboratory and bred together to generate *Cxcr4*<sup>+/-</sup> mice on a homozygous CD45.2 background. *Cxcr4*<sup>+/+</sup> mice were obtained from Jackson Laboratory on the homozygous CD45.1 and homozygous CD45.2 backgrounds. Creation of WHIM knockin mice bearing a heterozygous *Cxcr4*<sup>S338X</sup> mutation has been previously described (Balabanian et al., 2012), and sperm from these mice were used for *in vitro* fertilization at NIH to impregnate female C57BL/6 mice from Taconic Farms to produce *Cxcr4*<sup>+/S338X</sup> on a homozygous CD45.2 background. Both *Cxcr4*<sup>+/-</sup> and *Cxcr4*<sup>+/S338X</sup> mice on the homozygous CD45.2 background were bred to *Cxcr4*<sup>+/+</sup> mice on a homozygous CD45.1 background to produce *Cxcr4*<sup>+/-</sup> and *Cxcr4*<sup>+/S338X</sup> mice on a heterozygous CD45.1/CD45.2 background. In this way, differentially marked donor bone marrow cells could be tracked *in vivo* during competitive transplantation experiments. Five million bone marrow cells (~50% from each donor) were transferred via tail vein injection into recipient mice that had undergone lethal irradiation (900 rads) 8 hr prior to transplant. The mice were fed with neomycin water for 4 weeks post-irradiation. All animal experiments were performed using an NIAID Animal Care and Use Committee (ACUC) approved protocol in approved and certified facilities.

For additional details, see Extended Experimental Procedures.

### ACCESSION NUMBERS

The database of Genotypes and Phenotypes (dbGaP) accession number for the sequence data reported in this paper is phs000856.v1.p1.

### SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, seven figures, and two tables and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2015.01.014>.

### AUTHOR CONTRIBUTIONS

Patient recruitment and care was provided by D.H.M., D.V., M.M.M., N.K., S.P., M.A.R., K.R.C., I.M., R.D., and P.M.M. Experimental design was provided by D.H.M., J.G., Q.L., and P.M.M. Generation and analysis of experimental data were provided by J.G., Q.L., M.S., C.M., P.J., C.R.B., Z.D., E.S., E.Y., N.T., and D.A.L.P. with supervision and analysis by D.H.M., S.F.P., H.L.M., K.L.H., D.B.K., and P.M.M. K.B. and F.B. provided the WHIM knockin mice. D.H.M., J.G., and P.M.M. were principally responsible for writing the manuscript.

## ACKNOWLEDGMENTS

This work was supported by the Division of Intramural Research (DIR) of the National Institute of Allergy and Infectious Diseases (NIAID), NIH and was funded in part with federal funds from the National Cancer Institute, NIH under contract HHSN261200800001E. We thank Hongwei Zhang for flow cytometry advice, Carolyn Boscia for help with graphics, and the research subjects for their participation in the study. A provisional patent on CXCR4 knockdown as a method to enhance HSC engraftment has been filed by the US government with D.H.M., Q.L., M.S., J.G., H.L.M., and P.M.M. as inventors.

Received: August 22, 2014

Revised: November 3, 2014

Accepted: January 5, 2015

Published: February 5, 2015

## REFERENCES

Al Ustwani, O., Kurzrock, R., and Wetzler, M. (2014). Genetics on a WHIM. *Br. J. Haematol.* 164, 15–23.

Bachelerie, F., Ben-Baruch, A., Burkhardt, A.M., Combadiere, C., Farber, J.M., Graham, G.J., Horuk, R., Sparre-Ulrich, A.H., Locati, M., Luster, A.D., et al. (2014). International Union of Basic and Clinical Pharmacology. LXXXIX. Update on the extended family of chemokine receptors and introducing a new nomenclature for atypical chemokine receptors. *Pharmacol. Rev.* 66, 1–79.

Balabanian, K., Brotin, E., Blajoux, V., Bouchet-Delbos, L., Lainey, E., Fenechteau, O., Bonnet, D., Fiette, L., Emilie, D., and Bachelerie, F. (2012). Proper desensitization of CXCR4 is required for lymphocyte development and peripheral compartmentalization in mice. *Blood* 119, 5722–5730.

Beaussant Cohen, S., Fenneteau, O., Plouvier, E., Rohrlich, P.S., Daltroff, G., Plantier, I., Dupuy, A., Kerob, D., Beaupain, B., Bordigoni, P., et al. (2012). Description and outcome of a cohort of 8 patients with WHIM syndrome from the French Severe Chronic Neutropenia Registry. *Orphanet J. Rare Dis.* 7, 71.

Beul, C.C., Fuhlbrigge, R.C., Casasnovas, J.M., Aiuti, A., and Springer, T.A. (1996). A highly efficacious lymphocyte chemoattractant, stromal cell-derived factor 1 (SDF-1). *J. Exp. Med.* 184, 1101–1109.

Broxmeyer, H.E., Cooper, S., Kohli, L., Hangoc, G., Lee, Y., Mantel, C., Clapp, D.W., and Kim, C.H. (2003a). Transgenic expression of stromal cell-derived factor-1/CXC chemokine ligand 12 enhances myeloid progenitor cell survival/antiapoptosis in vitro in response to growth factor withdrawal and enhances myelopoiesis in vivo. *J. Immunol.* 170, 421–429.

Broxmeyer, H.E., Kohli, L., Kim, C.H., Lee, Y., Mantel, C., Cooper, S., Hangoc, G., Shaheen, M., Li, X., and Clapp, D.W. (2003b). Stromal cell-derived factor-1/CXCL12 directly enhances survival/antiapoptosis of myeloid progenitor cells through CXCR4 and G(α $\alpha$ ) proteins and enhances engraftment of competitive, repopulating stem cells. *J. Leukoc. Biol.* 73, 630–638.

Broxmeyer, H.E., Orschell, C.M., Clapp, D.W., Hangoc, G., Cooper, S., Plett, P.A., Liles, W.C., Li, X., Graham-Evans, B., Campbell, T.B., et al. (2005). Rapid mobilization of murine and human hematopoietic stem and progenitor cells with AMD3100, a CXCR4 antagonist. *J. Exp. Med.* 207, 1307–1318.

Challen, G.A., Sun, D., Jeong, M., Luo, M., Jelinek, J., Berg, J.S., Bock, C., Vasanthakumar, A., Gu, H., Xi, Y., et al. (2012). Dnmt3a is essential for hematopoietic stem cell differentiation. *Nat. Genet.* 44, 23–31.

Dale, D.C., Bolyard, A.A., Kelley, M.L., Westrup, E.C., Makaryan, V., Aprikyan, A., Wood, B., and Hsu, F.J. (2011). The CXCR4 antagonist plerixafor is a potential therapy for myelokathexis, WHIM syndrome. *Blood* 118, 4963–4966.

Dar, A., Kollet, O., and Lapidot, T. (2006). Mutual, reciprocal SDF-1/CXCR4 interactions between hematopoietic and bone marrow stromal cells regulate human stem cell migration and development in NOD/SCID chimeric mice. *Exp. Hematol.* 34, 967–975.

Dotta, L., Tassone, L., and Badolato, R. (2011). Clinical and genetic features of Warts, Hypogammaglobulinemia, Infections and Myelokathexis (WHIM) syndrome. *Curr. Mol. Med.* 11, 317–325.

Haribabu, B., Richardson, R.M., Fisher, I., Sozzani, S., Peiper, S.C., Horuk, R., Ali, H., and Snyderman, R. (1997). Regulation of human chemokine receptors CXCR4. Role of phosphorylation in desensitization and internalization. *J. Biol. Chem.* 272, 28726–28731.

Hernandez, P.A., Gorlin, R.J., Lukens, J.N., Taniuchi, S., Bohinjec, J., Francois, F., Klotman, M.E., and Diaz, G.A. (2003). Mutations in the chemokine receptor gene CXCR4 are associated with WHIM syndrome, a combined immunodeficiency disease. *Nat. Genet.* 34, 70–74.

Hirschhorn, R. (2003). In vivo reversion to normal of inherited mutations in humans. *J. Med. Genet.* 40, 721–728.

Jones, M.J., and Jallepalli, P.V. (2012). Chromothripsis: chromosomes in crisis. *Dev. Cell* 23, 908–917.

Kawai, T., and Malech, H.L. (2009). WHIM syndrome: congenital immune deficiency disease. *Curr. Opin. Hematol.* 16, 20–26.

Kawai, T., Choi, U., Cardwell, L., DeRavin, S.S., Naumann, N., Whiting-Theobald, N.L., Linton, G.F., Moon, J., Murphy, P.M., and Malech, H.L. (2007). WHIM syndrome myelokathexis reproduced in the NOD/SCID mouse xenotransplant model engrafted with healthy human stem cells transduced with C-terminus-truncated CXCR4. *Blood* 109, 78–84.

Kloosterman, W.P., Guryev, V., van Roosmalen, M., Duran, K.J., de Bruijn, E., Bakker, S.C., Letteboer, T., van Nesselrooij, B., Hochstenbach, R., Poot, M., and Cuppen, E. (2011). Chromothripsis as a mechanism driving complex de novo structural rearrangements in the germline. *Hum. Mol. Genet.* 20, 1916–1924.

Korbel, J.O., and Campbell, P.J. (2013). Criteria for inference of chromothripsis in cancer genomes. *Cell* 152, 1226–1236.

Krill, C.E., Jr., Smith, H.D., and Mauer, A.M. (1964). Chronic idiopathic granulocytopenia. *N. Engl. J. Med.* 270, 973–979.

Laurenti, E., Varnum-Finney, B., Wilson, A., Ferrero, I., Blanco-Bose, W.E., Ehninger, A., Knoepfler, P.S., Cheng, P.F., MacDonald, H.R., Eisenman, R.N., et al. (2008). Hematopoietic stem cell function and survival depend on c-Myc and N-Myc activity. *Cell Stem Cell* 3, 611–624.

McDermott, D.H., Liu, Q., Ulrick, J., Kwatemaa, N., Anaya-O'Brien, S., Penzak, S.R., Filho, J.O., Priel, D.A., Kelly, C., Garofalo, M., et al. (2011). The CXCR4 antagonist plerixafor corrects panleukopenia in patients with WHIM syndrome. *Blood* 118, 4957–4962.

McDermott, D.H., Liu, Q., Velez, D., Lopez, L., Anaya-O'Brien, S., Ulrick, J., Kwatemaa, N., Starling, J., Fleisher, T.A., Priel, D.A., et al. (2014). A phase 1 clinical trial of long-term, low-dose treatment of WHIM syndrome with the CXCR4 antagonist plerixafor. *Blood* 123, 2308–2316.

Nie, Y., Han, Y.C., and Zou, Y.R. (2008). CXCR4 is required for the quiescence of primitive hematopoietic cells. *J. Exp. Med.* 205, 777–783.

Orelio, C., Peeters, M., Haak, E., van der Horn, K., and Dzierzak, E. (2009). Interleukin-1 regulates hematopoietic progenitor and stem cells in the midgestation mouse fetal liver. *Haematologica* 94, 462–469.

Signoret, N., Rosenkilde, M.M., Klasse, P.J., Schwartz, T.W., Malim, M.H., Hoxie, J.A., and Marsh, M. (1998). Differential regulation of CXCR4 and CCR5 endocytosis. *J. Cell Sci.* 111, 2819–2830.

Stephens, P.J., Greenman, C.D., Fu, B., Yang, F., Bignell, G.R., Mudie, L.J., Pleasance, E.D., Lau, K.W., Beare, D., Stebbings, L.A., et al. (2011). Massive genomic rearrangement acquired in a single catastrophic event during cancer development. *Cell* 144, 27–40.

Sugiyama, T., Kohara, H., Noda, M., and Nagasawa, T. (2006). Maintenance of the hematopoietic stem cell pool by CXCL12-CXCR4 chemokine signaling in bone marrow stromal cell niches. *Immunity* 25, 977–988.

Tassone, L., Notarangelo, L.D., Bonomi, V., Savoldi, G., Sensi, A., Sorensina, A., Smith, C.I., Porta, F., Plebani, A., and Badolato, R. (2009). Clinical and genetic diagnosis of warts, hypogammaglobulinemia, infections, and

myelokathexis syndrome in 10 patients. *J Allergy Clin. Immunol.* 123, 1170–1173.

Venkatesan, S., Rose, J.J., Lodge, R., Murphy, P.M., and Foley, J.F. (2003). Distinct mechanisms of agonist-induced endocytosis for human chemokine receptors CCR5 and CXCR4. *Mol. Biol. Cell* 14, 3305–3324.

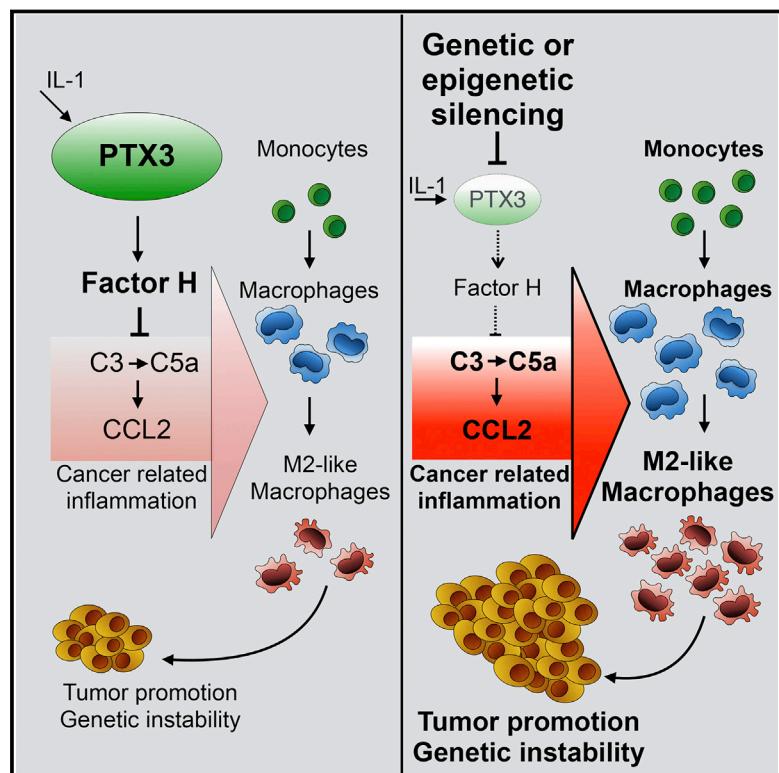
Wetzler, M., Talpaz, M., Kleinerman, E.S., King, A., Huh, Y.O., Guterman, J.U., and Kurzrock, R. (1990). A new familial immunodeficiency disorder characterized by severe neutropenia, a defective marrow release mechanism, and hypogammaglobulinemia. *Am. J. Med.* 89, 663–672.

Zou, Y.R., Kottmann, A.H., Kuroda, M., Taniuchi, I., and Littman, D.R. (1998). Function of the chemokine receptor CXCR4 in haematopoiesis and in cerebellar development. *Nature* 393, 595–599.

Zuelzer, W.W. (1964). “Myelokathexis”—a New Form of Chronic Granulocytopenia. Report of a Case. *N. Engl. J. Med.* 270, 699–704.

# PTX3 Is an Extrinsic Oncosuppressor Regulating Complement-Dependent Inflammation in Cancer

## Graphical Abstract



## Authors

Eduardo Bonavita, Stefania Gentile, ..., Cecilia Garlanda, Alberto Mantovani

## Correspondence

cecilia.garlanda@humanitasresearch.it  
(C.G.),  
alberto.mantovani@humanitasresearch.it  
(A.M.)

## In Brief

PTX3 deficiency triggers Complement-dependent tumor-promoting inflammation, with enhanced tumor burden, macrophage infiltration, cytokine production, angiogenesis, and genetic instability, revealing the role of this innate immunity mediator as an extrinsic oncosuppressor.

## Highlights

- PTX3 deficiency unleashes Complement-dependent tumor-promoting inflammation
- Tumors developed in a PTX3-deficient context have higher frequency of mutated *Trp53*
- PTX3 expression is epigenetically repressed in selected human tumors
- Complement is an essential component of tumor-promoting inflammation

# PTX3 Is an Extrinsic Oncosuppressor Regulating Complement-Dependent Inflammation in Cancer

Eduardo Bonavita,<sup>1</sup> Stefania Gentile,<sup>1</sup> Marcello Rubino,<sup>1</sup> Virginia Maina,<sup>1</sup> Roberto Papait,<sup>1,2</sup> Paolo Kunderfranco,<sup>1</sup> Carolina Greco,<sup>1</sup> Francesca Feruglio,<sup>1</sup> Martina Molgora,<sup>1</sup> Ilaria Laface,<sup>1</sup> Silvia Tartari,<sup>1</sup> Andrea Doni,<sup>1</sup> Fabio Pasqualini,<sup>1</sup> Elisa Barbat,<sup>1</sup> Gianluca Basso,<sup>1</sup> Maria Rosaria Galdiero,<sup>1</sup> Manuela Nebuloni,<sup>3</sup> Massimo Roncalli,<sup>1</sup> Piergiuseppe Colombo,<sup>1</sup> Luigi Laghi,<sup>1</sup> John D. Lambris,<sup>4</sup> Sébastien Jaillon,<sup>1</sup> Cecilia Garlanda,<sup>1,\*</sup> and Alberto Mantovani<sup>1,5,\*</sup>

<sup>1</sup>Humanitas Clinical and Research Center, Rozzano (Milan) 20089, Italy

<sup>2</sup>Institute of Genetics and Biomedical Research, National Research Council, Rozzano (Milan) 20089, Italy

<sup>3</sup>Pathology Unit, L. Sacco Department of Clinical Sciences, University of Milan, Milan 20157, Italy

<sup>4</sup>Department of Pathology and Laboratory Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

<sup>5</sup>Humanitas University, Rozzano (Milan) 20089, Italy

\*Correspondence: cecilia.garlanda@humanitasresearch.it (C.G.), alberto.mantovani@humanitasresearch.it (A.M.)

<http://dx.doi.org/10.1016/j.cell.2015.01.004>

## SUMMARY

PTX3 is an essential component of the humoral arm of innate immunity, playing a nonredundant role in resistance against selected microbes and in the regulation of inflammation. PTX3 activates and regulates the Complement cascade by interacting with C1q and with Factor H. PTX3 deficiency was associated with increased susceptibility to mesenchymal and epithelial carcinogenesis. Increased susceptibility of *Ptx3*<sup>-/-</sup> mice was associated with enhanced macrophage infiltration, cytokine production, angiogenesis, and *Trp53* mutations. Correlative evidence, gene-targeted mice, and pharmacological blocking experiments indicated that PTX3 deficiency resulted in amplification of Complement activation, CCL2 production, and tumor-promoting macrophage recruitment. PTX3 expression was epigenetically regulated in selected human tumors (e.g., leiomyosarcomas and colorectal cancer) by methylation of the promoter region and of a putative enhancer. Thus, PTX3, an effector molecule belonging to the humoral arm of innate immunity, acts as an extrinsic oncosuppressor gene in mouse and man by regulating Complement-dependent, macrophage-sustained, tumor-promoting inflammation.

## INTRODUCTION

Inflammatory cells and molecules are an essential component of the tumor microenvironment (Coussens et al., 2013; Grivennikov et al., 2010; Hanahan and Weinberg, 2011; Mantovani et al., 2008). The connection between inflammation and cancer can be viewed as consisting of two pathways. Selected inflammatory conditions, the prototype of which is ulcerative colitis, increase the risk of developing cancer (extrinsic pathway). On the other hand genetic events that cause cancer orchestrate the construction of an inflammatory microenvironment even in tumors that are

epidemiologically unrelated to inflammation (Mantovani et al., 2008). Genetic approaches have unequivocally proven the role of cells and mediators (e.g., cytokines) of the innate and adaptive immune system in sustaining tumor-promoting inflammation. However, no dominant or suppressor oncogene encodes for effector molecules of the humoral innate immune system.

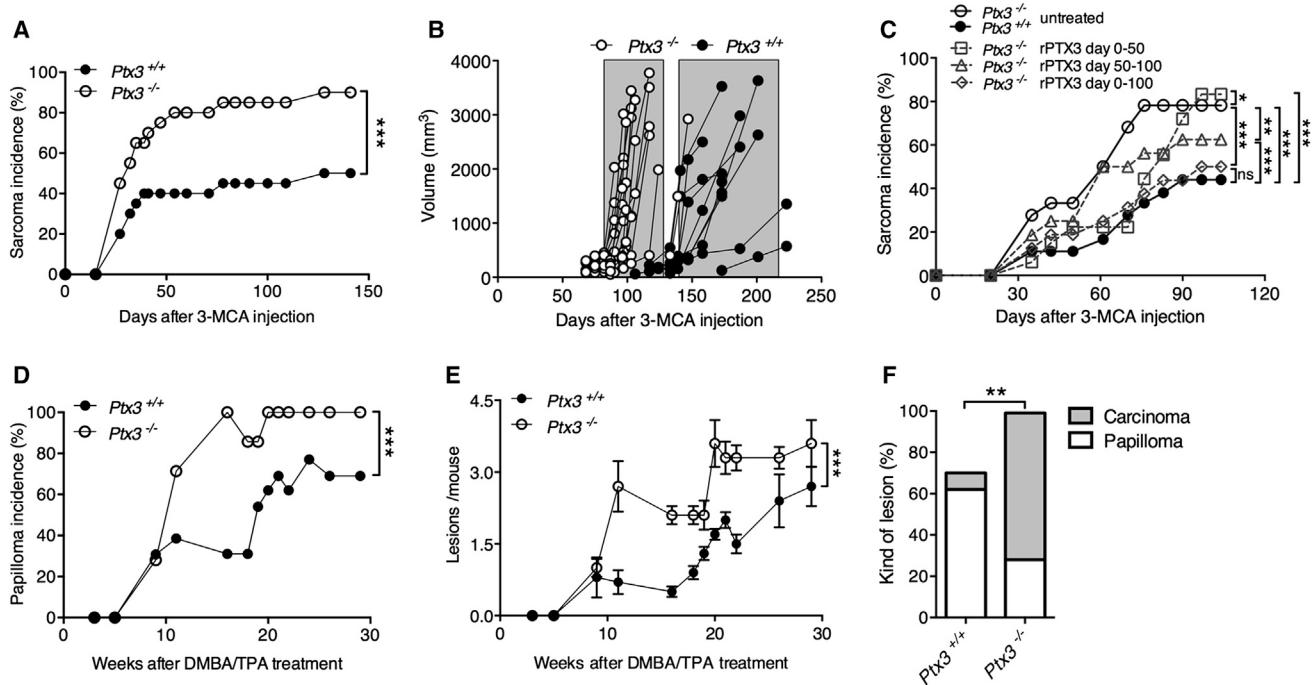
The innate immune system includes a cellular and a humoral arm (Garlanda et al., 2005). The humoral arm of innate immunity is constituted of diverse molecules including Complement components, collectins, ficolins, and pentraxins (Bottazzi et al., 2010). The long pentraxin PTX3 is an essential component of the humoral arm of innate immunity (Bottazzi et al., 2010; Garlanda et al., 2005). PTX3 acts as a functional ancestor of antibodies: it interacts with selected microbial molecules (Jeannin et al., 2005); it has opsonic activity via Fc $\gamma$  receptors (Moalli et al., 2010); it activates and regulates the Complement cascade by interacting with C1q and Factor H (Bottazzi et al., 1997; Deban et al., 2008); and it regulates inflammation by interacting with P-selectin via its glycosidic moiety (Deban et al., 2010). PTX3 plays a nonredundant role in resistance against selected microbial pathogens in mouse and man (Chiarini et al., 2010; Cunha et al., 2014; Garlanda et al., 2002) and is a candidate therapeutic agent against *Aspergillus fumigatus* and *Pseudomonas aeruginosa*.

In the context of an effort to translate this molecule to the clinic as a novel candidate antimicrobial agent in cancer patients at risk of infection (Cunha et al., 2014), it was important to assess its role in carcinogenesis. We found that the humoral pattern recognition molecule PTX3 acts as an oncosuppressor in mice and humans by regulating Complement-dependent tumor-promoting inflammation. This observation represents a missing link in the connection between inflammation and cancer by providing genetic evidence that an effector molecule of humoral innate immunity can act as a cancer gene.

## RESULTS

### PTX3 Deficiency Increases Susceptibility to Cancer Development and Growth

In order to address the role of PTX3 in cancer-related inflammation (CRI) and tumor development, we analyzed the susceptibility



**Figure 1. PTX3 Deficiency Increases Susceptibility to Carcinogenesis**

(A and B) Incidence (A) and tumor volume (B) of 3-MCA-induced sarcoma in *Ptx3*<sup>-/-</sup> and *Ptx3*<sup>+/+</sup> mice. In (B), the volume of each tumor is shown. One experiment with 20 mice/group out of ten performed with similar results is shown.

(C) Effect of recombinant PTX3 (9 µg/day, s.c. by osmotic pump) administered from day 0 to 50,–100, or from 0–100 on tumor incidence.

(D and E) Incidence (D) and number of lesions (mean ± SEM) (E) of DMBA/TPA-induced skin papillomas in *Ptx3*<sup>-/-</sup> and *Ptx3*<sup>+/+</sup> mice. One experiment with ten mice/group out of five performed with similar results is shown.

(F) Histological classification of lesions developed in *Ptx3*<sup>-/-</sup> and *Ptx3*<sup>+/+</sup> mice. \*\*p < 0.01, Chi-square test. Two pooled experiments.

(A–E) \*p < 0.05, \*\*p < 0.01, \*\*\*p < 0.001 paired Student's t test. See also Figure S1.

of *Ptx3*<sup>-/-</sup> mice in different models of mesenchymal or epithelial carcinogenesis. As shown in Figure 1A, in the model of 3-Methylcholanthrene (3-MCA)-induced carcinogenesis, 80% *Ptx3*<sup>-/-</sup> mice developed a sarcoma in comparison to 40% *Ptx3*<sup>+/+</sup> mice ( $p = 0.0002$ ). *Ptx3*<sup>-/-</sup> sarcomas grew faster than *Ptx3*<sup>+/+</sup> tumors and reached the endpoint (3 cm<sup>3</sup>) in 100–120 days, whereas *Ptx3*<sup>+/+</sup> tumors reached it in 170–220 days (Figure 1B). Treatment with recombinant PTX3 subcutaneously (s.c.) throughout the carcinogenesis process (days 0–100) completely rescued the phenotype of *Ptx3*<sup>-/-</sup> mice (Figure 1C). Early treatment (days 0–50) inhibited the early difference, with an increase in incidence following interruption of treatment. A minor decrease in tumor incidence was also observed when treatment was begun on day 50 through day 100 (Figure 1C). The impact of PTX3 administration after day 50 may well reflect a role of Complement (see below) in promotion of established transplanted tumors (Markiewski et al., 2008). Treatment of *Ptx3*<sup>+/+</sup> mice caused some delay in tumor appearance but did not affect incidence (Figure S1A available online).

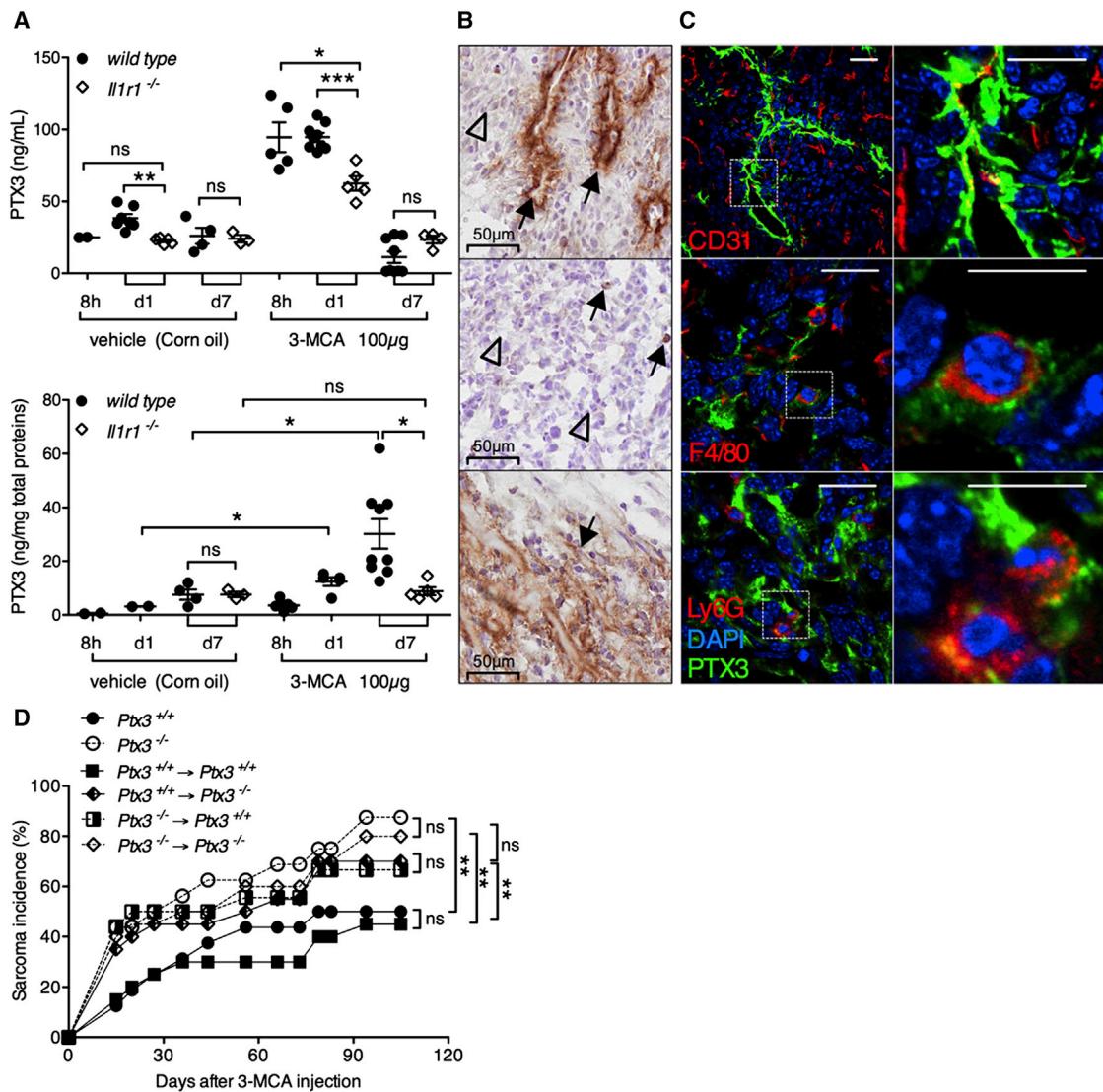
We next addressed the relevance of PTX3 deficiency in a model of 7,12-dimethylbenz [α] anthracene/terephthalic acid (DMBA/TPA)-induced skin carcinogenesis. As shown in Figures 1D and 1E, both the incidence and the multiplicity of papillomas were significantly increased in *Ptx3*<sup>-/-</sup> mice in comparison to *Ptx3*<sup>+/+</sup> mice ( $p = 0.0004$  and  $p = 0.0005$ , respectively). In addition,

the number of lesions evolving to skin carcinomas was significantly higher in *Ptx3*<sup>-/-</sup> mice ( $p = 0.009$ ), suggesting a more aggressive behavior of DMBA/TPA-induced skin lesion in PTX3-deficient hosts (Figure 1F).

These data indicate that PTX3 is involved in controlling the incidence and growth of tumors of both mesenchymal and epithelial origin.

#### Production of PTX3 during Carcinogenesis

The following experiments, aimed at defining the underlying cellular and molecular mechanisms, were focused on the 3-MCA-induced carcinogenesis model. We first analyzed the production of PTX3 during carcinogenesis. 3-MCA is a carcinogenic compound acting as initiator and promoter, metabolized by the liver. PTX3 levels were increased in serum (Figure 2A) and liver (not shown) a few hours (8 and 24) after 3-MCA treatment, whereas after 7 days they had returned to basal levels. At the injection site, the levels of PTX3 were higher 24 hr after 3-MCA treatment in comparison to vehicle and further increased at day 7. PTX3 was originally cloned as an IL-1 inducible gene (Garlanda et al., 2005). It was therefore important to assess whether IL-1 was upstream of PTX3 in 3-MCA carcinogenesis. Both systemic and local PTX3 levels were significantly lower in *Il1r1*<sup>-/-</sup> mice compared to wild-type mice, suggesting that in this model, IL-1 is involved in PTX3 induction (Figure 2A). In established



**Figure 2. Analysis of the Cellular Source of PTX3 and of Its Role in 3-MCA-Induced Carcinogenesis**

(A) Systemic (upper panel) and local (lower panel) PTX3 production at different time points upon vehicle or 3-MCA injection in wild-type and *II1r1<sup>-/-</sup>* mice (mean ± SEM). \*p < 0.05, \*\*p < 0.01, \*\*\*p < 0.001, unpaired Student's t test.

(B) Immunohistochemical analysis of PTX3 expression in 3-MCA-induced sarcoma. Arrows show vessels (upper panel), infiltrating leukocytes (central panel), and interstitial stroma (lower panel). Arrowheads show tumor cells (upper and central panels).

(C) Immunofluorescence analysis of PTX3 expression in blood vessels (CD31<sup>+</sup> cells), macrophages (F4/80<sup>+</sup> cells), and neutrophils (Ly6G<sup>+</sup> cells). Scale bar represents 150 µm.

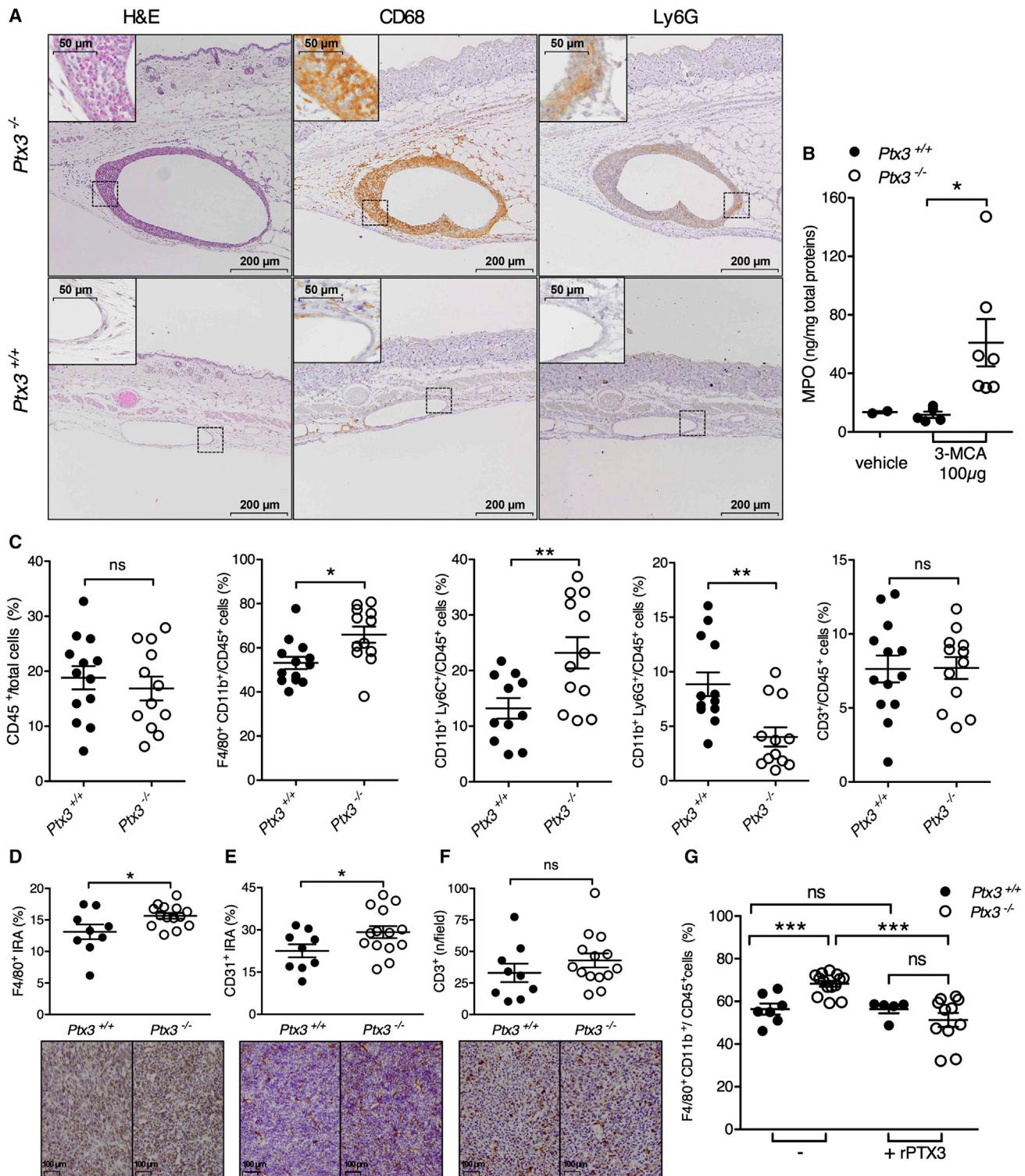
(D) 3-MCA-induced sarcoma incidence in PTX3-chimeric mice (Donor → Recipient). \*\*p < 0.01, paired Student's t test.

See also Figure S2.

tumors (day 100), PTX3 immunostaining was associated with vessels, macrophages, neutrophils, and interstitial stroma, but not to tumor cells (Figures 2B and 2C). Accordingly, most sarcoma cell lines isolated from *Ptx3<sup>+/+</sup>* tumor-bearing mice did not produce PTX3 in basal or inflammatory conditions (Figure S2C). PTX3 expression in cells of the monocyte-macrophage lineage was confirmed by RT-PCR analysis, with highest levels in MCHII<sup>low</sup> macrophages (see below) (Figure S2D). At early time points (e.g., day 15), PTX3 expression was strongly induced in F4/80<sup>+</sup> cells (macrophages) in the skin and to a lesser extent in

liver, as well as in stromal cells (Figure S2A). In the latter cell type, substantial coexpression of a marker of DNA damage response (DDR), γ-H2Ax, was observed only at early time points to wane thereafter (Figure S2B). Collectively these results suggest that infiltrating leukocytes and endothelial cells are a major source of PTX3 and that silencing of PTX3 expression in tumor cells occurs generally during progression.

Bone marrow chimeras were used to assess the relative importance of hematopoietic and nonhematopoietic (including tumor) cells as a source of PTX3 and in the protection against

**Figure 3. PTX3 Deficiency Is Associated to Increased CRI**

(A) Histological and immunohistochemical analysis of macrophage (CD68<sup>+</sup>) and neutrophil (Ly6G<sup>+</sup>) infiltration in *Ptx3*<sup>-/-</sup> and *Ptx3*<sup>+/+</sup> 3-MCA-injection site at 7 days.

(B) MPO levels in 3-MCA-injection site at 7 days (mean  $\pm$  SEM). See also Table S1.

(C) Analysis of the leukocyte infiltrate in 3-MCA-sarcoma by FACS. Total leukocytes (CD45<sup>+</sup>), macrophages (F4/80<sup>+</sup>/CD11b<sup>+</sup>), monocytes (CD11b<sup>+</sup>/Ly6C<sup>+</sup>), neutrophils (CD11b<sup>+</sup>/Ly6G<sup>+</sup>), and lymphocytes (CD3<sup>+</sup>) were analyzed (mean  $\pm$  SEM).

(legend continued on next page)

carcinogenesis. PTX3 tumor concentration in the two chimeric groups (*Ptx3*<sup>+/+</sup> mice receiving *Ptx3*<sup>-/-</sup> bone marrow and *Ptx3*<sup>-/-</sup> mice receiving *Ptx3*<sup>+/+</sup> bone marrow) were comparable (2.15 ng/mg ± 0.8 ng/mg and 1.89 ng/mg ± 0.1 ng/mg) and significantly lower than in *Ptx3*<sup>+/+</sup> tumors (6.61 ng/mg ± 0.2 ng/mg, p < 0.003), suggesting that both nonhematopoietic cells and hematopoietic cells produce PTX3 in tumors. PTX3 competence in hematopoietic cells or in stromal cells was insufficient to reconstitute the full “protected” phenotype observed in *Ptx3*<sup>+/+</sup> mice or in reconstituted fully competent mice (Figure 2D). Thus, PTX3 derived from both hematopoietic and nonhematopoietic cells contributes to protection against carcinogenesis.

### PTX3 Deficiency Is Associated with Increased Cancer-Related Inflammation

PTX3 does not affect the proliferation of normal and transformed cells (Figure S1B). It interacts with Fc $\gamma$  receptors (Lu et al., 2008; Moalli et al., 2010) and has antibody-like properties (Bottazzi et al., 2010) including the capacity to regulate inflammatory reactions. Therefore, we focused on CRI in *Ptx3*<sup>-/-</sup> mice. We first evaluated leukocyte recruitment into the 3-MCA-injection site at early time points in *Ptx3*<sup>+/+</sup> and *Ptx3*<sup>-/-</sup> mice. As shown in Figure 3A, on day 7 after injection macrophages (CD68 $^{+}$  cells) and neutrophils (Ly6G $^{+}$  cells) recruited around 3-MCA-containing corn oil droplets were more abundant in *Ptx3*<sup>-/-</sup> compared to *Ptx3*<sup>+/+</sup> mice. In agreement, the MPO content (Figure 3B), as well as the levels of CCL2 and CXCL2 (Table S1), were significantly higher in homogenates of tissues collected from *Ptx3*<sup>-/-</sup> mice, compared to *Ptx3*<sup>+/+</sup> mice.

We next analyzed the leukocyte infiltrate of tumors collected at their maximum volume (3 cm<sup>3</sup>) by flow cytometry. As shown in Figure 3C, the percentage of macrophages (F4/80 $^{+}$ /CD11b $^{+}$  cells) and monocytes (Ly6C $^{\text{high}}$ /CD11b $^{+}$  cells) in CD45 $^{+}$  cells was higher in *Ptx3*<sup>-/-</sup> compared to *Ptx3*<sup>+/+</sup> mice, whereas the percentage of neutrophils (Ly6G $^{\text{high}}$ /CD11b $^{+}$  cells) was significantly lower. The percentage of T lymphocytes (CD3 $^{+}$  cells) was comparable in the two groups. Similar results were obtained when the analysis was performed at day 100 after 3-MCA-treatment in tumors of different size in *Ptx3*<sup>-/-</sup> and *Ptx3*<sup>+/+</sup> mice (Figure S3A). Histological analysis confirmed increased macrophage infiltration in *Ptx3*<sup>-/-</sup> sarcomas and showed higher density of CD31 $^{+}$  vessels (Figures 3D and 3E). Similar lymphocyte infiltration in the two groups was confirmed by histological analysis (Figure 3F).

In addition, the concentrations of proinflammatory cytokines TNF $\alpha$ , IL-1 $\beta$  and IL-6, CCL2, and the proangiogenic vascular endothelial growth factor (VEGF) were significantly higher in *Ptx3*<sup>-/-</sup> tumor homogenates than in *Ptx3*<sup>+/+</sup>. IL-6 and VEGF levels were significantly higher also at the systemic level in *Ptx3*<sup>-/-</sup> mice (Table S1). Treatment with recombinant PTX3 reduced macrophage infiltration and angiogenesis abolishing the difference between the two genotypes (Figures 3G and S3B).

### PTX3 Deficiency Is Associated to Increased Complement Activation

PTX3 has been shown to tune inflammatory responses by modulating P-selectin-dependent leukocyte recruitment and Complement activation (Deban et al., 2010; Inforzato et al., 2013). We first addressed the role of PTX3-P-selectin interaction in leukocyte recruitment and in tumor growth by evaluating sarcoma incidence in P-selectin-deficient and in PTX3/P-selectin-double-deficient mice. As shown in Figure S3C, we found that PTX3 deficiency was associated to higher tumor incidence regardless of P-selectin competence. Moreover, P-selectin deficiency did not affect susceptibility to 3-MCA carcinogenesis.

Next, we analyzed Complement activation. Confocal microscopy of the injection site revealed significantly higher C3 immunoreactivity in areas surrounding 3-MCA-containing oil droplets in *Ptx3*<sup>-/-</sup> mice, in comparison to *Ptx3*<sup>+/+</sup> mice, where C3 deposition was negligible (Figure 4A). In tumors, C3 deposition occurred also in *Ptx3*<sup>+/+</sup> lesions, but it was significantly lower than in *Ptx3*<sup>-/-</sup> sarcomas (Figure 4B, two pooled experiments). An analysis of C3 deposition and PTX3 expression in each *Ptx3*<sup>+/+</sup> lesion was performed on samples from one experiment and showed that C3 and PTX3 were significantly inversely correlated (Pearson r = -0.65, p = 0.03). Treatment with recombinant PTX3 reduced C3 immunoreactivity in *Ptx3*<sup>-/-</sup> tumors to levels observed in *Ptx3*<sup>+/+</sup> tumors (Figure 4C).

In addition, a higher C5a concentration was measured in *Ptx3*<sup>-/-</sup> tumor homogenates compared to *Ptx3*<sup>+/+</sup> ones (Figure 4D). In contrast, the immunoreactivity for C5b-9, which was associated with necrotic tumor areas, did not differ in *Ptx3*<sup>-/-</sup> and *Ptx3*<sup>+/+</sup> tumors (Figure 4E). We next analyzed the immunoreactivity for Factor H and C4BP, because PTX3 interacts with these two Complement-regulatory proteins without interfering with their functional activity and facilitates their deposition on damaged cells, amplifying their regulatory potential (Braunschweig and Józsi, 2011; Deban et al., 2008). As shown in Figure 4F, the immunostaining for Factor H in tumors from *Ptx3*<sup>-/-</sup> mice was significantly lower than in *Ptx3*<sup>+/+</sup> tumors. A correlation analysis showed that PTX3 expression and Factor H deposition were significantly positively correlated (Pearson r = 0.71, p = 0.01). We also observed lower C4BP deposition in *Ptx3*<sup>-/-</sup> than in *Ptx3*<sup>+/+</sup> tumors, but the difference did not reach statistical significance (p = 0.07) (Figure S4A).

### Role of Complement

In the effort to assess the pathogenic role of increased Complement activation associated to PTX3 deficiency, we evaluated the susceptibility to 3-MCA of C3-deficient and PTX3/C3-double-deficient mice. Carcinogenesis by 3-MCA and macrophage recruitment in C3<sup>-/-</sup> mice were significantly lower than in wild-type animals (Figures 4G and 4H). In addition, C3 deficiency reduced the susceptibility to 3-MCA of *Ptx3*<sup>-/-</sup> mice to levels comparable to that of *Ptx3*<sup>+/+</sup> mice and double deficiency was

(D–F) Analysis of macrophages (F4/80 $^{+}$  immunoreactive area, IRA), angiogenesis (CD31 $^{+}$  IRA), and lymphocytes (CD3 $^{+}$  cells) in 3-MCA-sarcoma by immunohistochemistry (each dot represents the average of ten fields, mean ± SEM).

(G) Analysis by FACS of the macrophage infiltrate (mean ± SEM) in 3-MCA-sarcoma of *Ptx3*<sup>-/-</sup> and *Ptx3*<sup>+/+</sup> mice, after treatment with recombinant PTX3 (from day 0 to 100). \*p < 0.05, \*\*p < 0.01, unpaired Student's t test.

See also Figure S3.

associated to macrophage recruitment comparable to that of  $C3^{-/-}$  mice. Finally, treatment with the C5a receptor antagonist PMX53 reduced the susceptibility of  $Ptx3^{-/-}$  mice, reverting the phenotype, without affecting the tumor growth in wild-type mice (Figure 4I).

Factor H-deficient mice are depleted in circulating C3 and suffer from spontaneous nephropathy (Pickering et al., 2002). Therefore, a genetic double deficiency approach would not be informative as to the actual role of Factor H. Thus, we set up an in vitro model of C3 deposition on  $Ptx3^{+/+}$  and  $Ptx3^{-/-}$  cell lines, in the presence of recombinant PTX3 and a recombinant Factor H inhibitor (SRC19-20) (Banda et al., 2013). For these experiments, we selected one wild-type sarcoma line (Wt-2), generated on day 100, which, unlike most wild-type cell lines (Figure S2C), produces detectable levels of PTX3 (8 ng/ml in 24 hr). In line with results obtained in vivo (Figure 4B), C3 deposition was significantly higher on  $Ptx3^{-/-}$  than on  $Ptx3^{+/+}$  producing cells (Figures 4J and 4K). Recombinant PTX3 reduced C3 deposition on  $Ptx3^{-/-}$  cells abrogating the difference between  $Ptx3^{+/+}$  and  $Ptx3^{-/-}$  cells, and SRC19-20 that interferes with the interaction between PTX3 and Factor H (Deban et al., 2008), caused higher C3 deposition on  $Ptx3^{+/+}$  cells and did not further increase C3 deposition on  $Ptx3^{-/-}$  cells. Notably, the rescue of C3 deposition on  $Ptx3^{-/-}$  cells obtained with recombinant PTX3 was abolished in the presence of SRC19-20.

To further address whether PTX3 regulates immune complexes (IC)-dependent Complement and leukocyte activation, we measured circulating IC and analyzed their composition and proinflammatory activity. PTX3 deficiency was associated with normal levels of total IC and C3-containing IC (Figure S4C). Qualitative analysis revealed the presence of PTX3 in IC, but the amount of C3, C1q, and total Ig did not differ between  $Ptx3^{+/+}$  and  $Ptx3^{-/-}$  IC (Figure S4D). Finally,  $Ptx3^{+/+}$  and  $Ptx3^{-/-}$  IC equally activated macrophages (expression of polarization markers and cytokines) indicating that PTX3 does not regulate CRI through IC in this cancer model (Figure S4E).

All together, these data suggest that Complement is an essential component of tumor-promoting inflammation and that the increased Complement activation associated to PTX3 deficiency due to lack of recruitment of Factor H has a major role in sustaining an exacerbated inflammatory response and enhanced carcinogenesis.

#### Other Components of Cancer-Related Inflammation

As shown in Figure 3C and Table S1, PTX3 deficiency was associated with increased macrophage infiltration and increased levels of the monocyte attracting chemokine CCL2. Complement activation and C5a in particular induce chemokine production. It was therefore important to assess the role of CCL2 in the increased susceptibility to carcinogenesis exhibited by PTX3-deficient mice. To this aim, we treated  $Ptx3^{-/-}$  and  $Ptx3^{+/+}$  mice with an anti-CCL2 Ab. As shown in Figure 5A, the anti-CCL2 treatment did not modify the tumor incidence of  $Ptx3^{+/+}$  mice, whereas it reduced the susceptibility of  $Ptx3^{-/-}$  mice, completely reverting the phenotype.

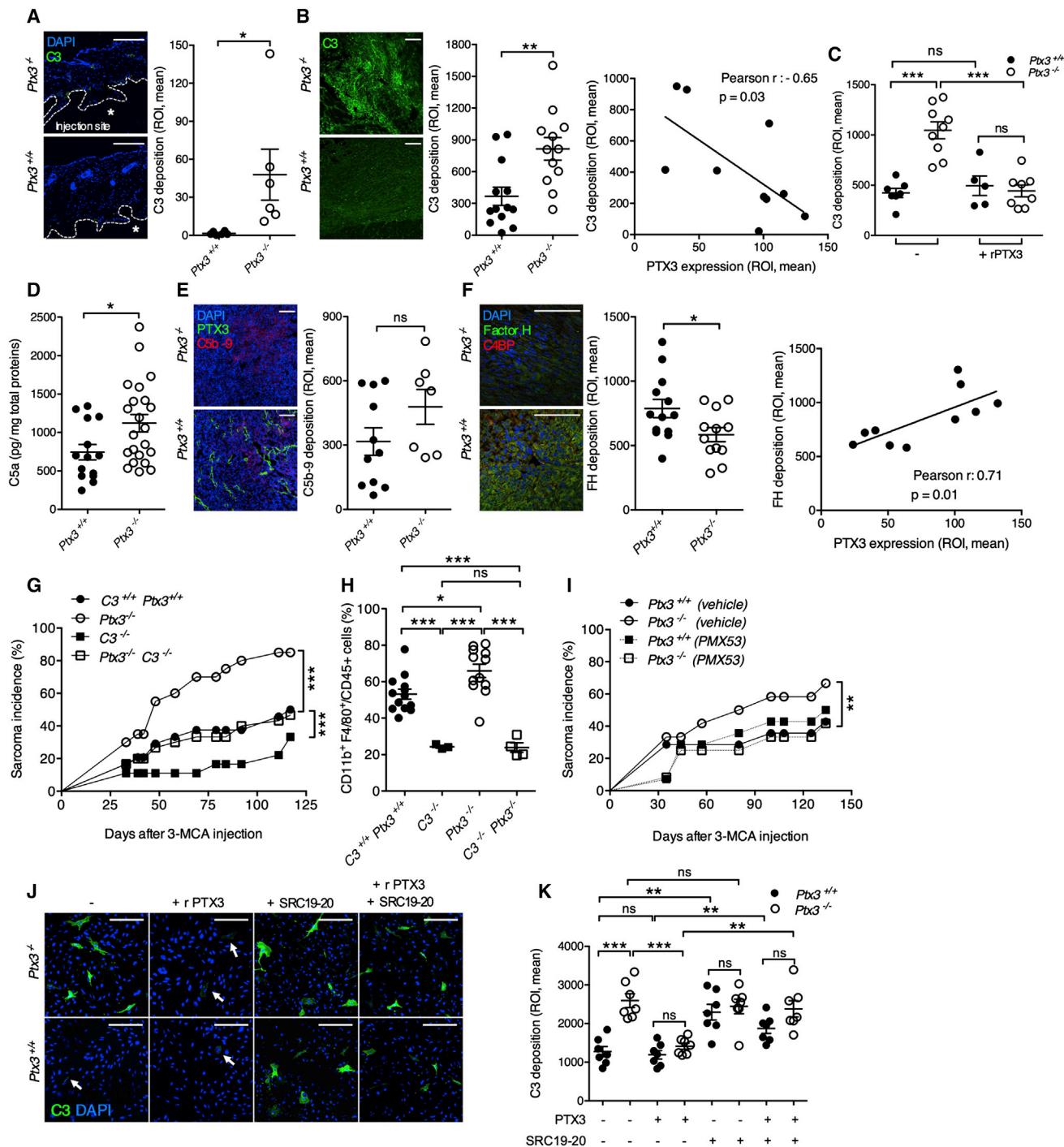
CCL2 has been shown to skew macrophage polarization in an M2-like direction (Roca et al., 2009; Sierra-Filardi et al., 2014), we therefore characterized the phenotype of tumor-associated

macrophages (TAM) in  $Ptx3^{-/-}$  and  $Ptx3^{+/+}$  mice. Upon purification of CD11b<sup>+</sup>/Ly6G<sup>-</sup> cells, MHCII<sup>high</sup> (CD11b<sup>+</sup>/Ly6C<sup>-</sup>/MHCII<sup>high</sup>) and MHCII<sup>low</sup> (CD11b<sup>+</sup>/Ly6C<sup>-</sup>/MHCII<sup>low</sup>) macrophages and monocytes (CD11b<sup>+</sup>/Ly6C<sup>+</sup>/MHCII<sup>-</sup>) were sorted to >98% purity to analyze the expression of M1- and M2-like markers as described (Laoui et al., 2014). As shown in Figures 5B–5D and Table S2,  $Ptx3^{-/-}$  MHCII<sup>high</sup> and MHCII<sup>low</sup> macrophages and monocytes expressed higher levels of M2-like markers (in particular Arg1, Ym1,Fizz1, Il10, and Mcr1 in macrophages and Ym1 and Fizz1 in monocytes) and lower levels of M1 markers (in particular Il12a and Ifng) in comparison to  $Ptx3^{+/+}$  macrophages and monocytes. Moreover  $Ptx3^{-/-}$  macrophages (MHCII<sup>high</sup> or MHCII<sup>low</sup>) expressed higher levels of Nos2, which is overexpressed by TAMs (Murray et al., 2014).

#### Increased *Trp53* Mutations and DNA Damage in $Ptx3^{-/-}$ Mice

Because CRI is potentially a cause of gene instability (Colotta et al., 2009), we next addressed whether the increased tumor incidence observed in  $Ptx3^{-/-}$  mice was associated to increased gene mutation. To this aim, we evaluated the frequency of mutations in *Trp53* and *Kras*, which are target genes mutated by 3-MCA, in sarcoma cell lines isolated from  $Ptx3^{-/-}$  and  $Ptx3^{+/+}$  tumor-bearing mice. The genomic DNA of 10  $Ptx3^{+/+}$  and 12  $Ptx3^{-/-}$  3-MCA-induced sarcomas was examined by direct sequencing of exons 5 to 8 of the *Trp53* gene and of exons 1 and 2 of *Kras* gene. In line with previous studies (Watanabe et al., 1999), eight out of ten wild-type sarcomas (80%) carried mutations of *Trp53* gene in 12 different codons, whereas *Kras* mutations were less frequent (20%) and were restricted to two codons (Table S3). As shown in Figure 6A, the number of *Trp53* mutations was significantly higher in  $Ptx3^{-/-}$  than in  $Ptx3^{+/+}$  sarcoma cell lines, whereas no differences were observed in *Kras* gene mutations (Figure 6B). Interestingly, similar numbers of mutations were observed in  $Ptx3^{-/-}$  cell lines collected 100 days upon treatment or at the end point (tumor volume 3 cm<sup>3</sup>) suggesting that the increased *Trp53* gene instability is not due to accumulation of mutations during tumor progression (Table S3). We finally addressed the functional consequences of *Trp53* mutations. As shown in Figure 6C, the expression of *Mdm2*, *Bax*, and *Cdkn1a* (p21) downstream p53 activation, was significantly reduced in  $Ptx3^{-/-}$  tumor cell lines in comparison to  $Ptx3^{+/+}$ , indicating that the higher incidence of *Trp53* mutations observed in  $Ptx3^{-/-}$  is likely associated with increased loss of onco-suppressor function.

To further investigate whether and to what extent the increased inflammation in  $Ptx3^{-/-}$  mice translated into DNA damage we studied DNA damage by performing immunohistochemical analysis of 8-OH-deoxiguanosine (8-OH-dG), a modified DNA base generated during oxidative damage. As shown in Figure 6D, the number of 8-OH-dG-positive nuclei was significantly higher in  $Ptx3^{-/-}$  than in  $Ptx3^{+/+}$  tissues 15 days after 3-MCA injection. In addition, the expression of the DDR markers,  $\gamma$ -H2Ax and 53BP1 was significantly higher in  $Ptx3^{-/-}$  tissues (Figures 6E and 6F). The relevance of damage caused by reactive species was addressed by treating 3-MCA-injected mice with amminiguaniidine hemisulphate (AG), an irreversible inhibitor of iNOS. AG significantly reduced the tumor incidence in  $Ptx3^{-/-}$  mice (Figure 6G). In addition, treatment with AG



**Figure 4. PTX3 Deficiency Is Associated to Increased Complement Deposition in 3-MCA-Sarcoma**

(A–C) Analysis by confocal microscopy of C3 deposition in the injection site (\*) at 7 days (A) or in tumors (B and C) of *Ptx3<sup>-/-</sup>* and *Ptx3<sup>+/+</sup>* mice, after treatment with recombinant PTX3 (from day 0 to 100) (C). Correlation analysis between C3 and PTX3 expression was performed in wild-type tumors (B).

(D) Analysis by ELISA of C5a concentrations, (mean  $\pm$  SEM, two pooled experiments), in 3-MCA-tumors.

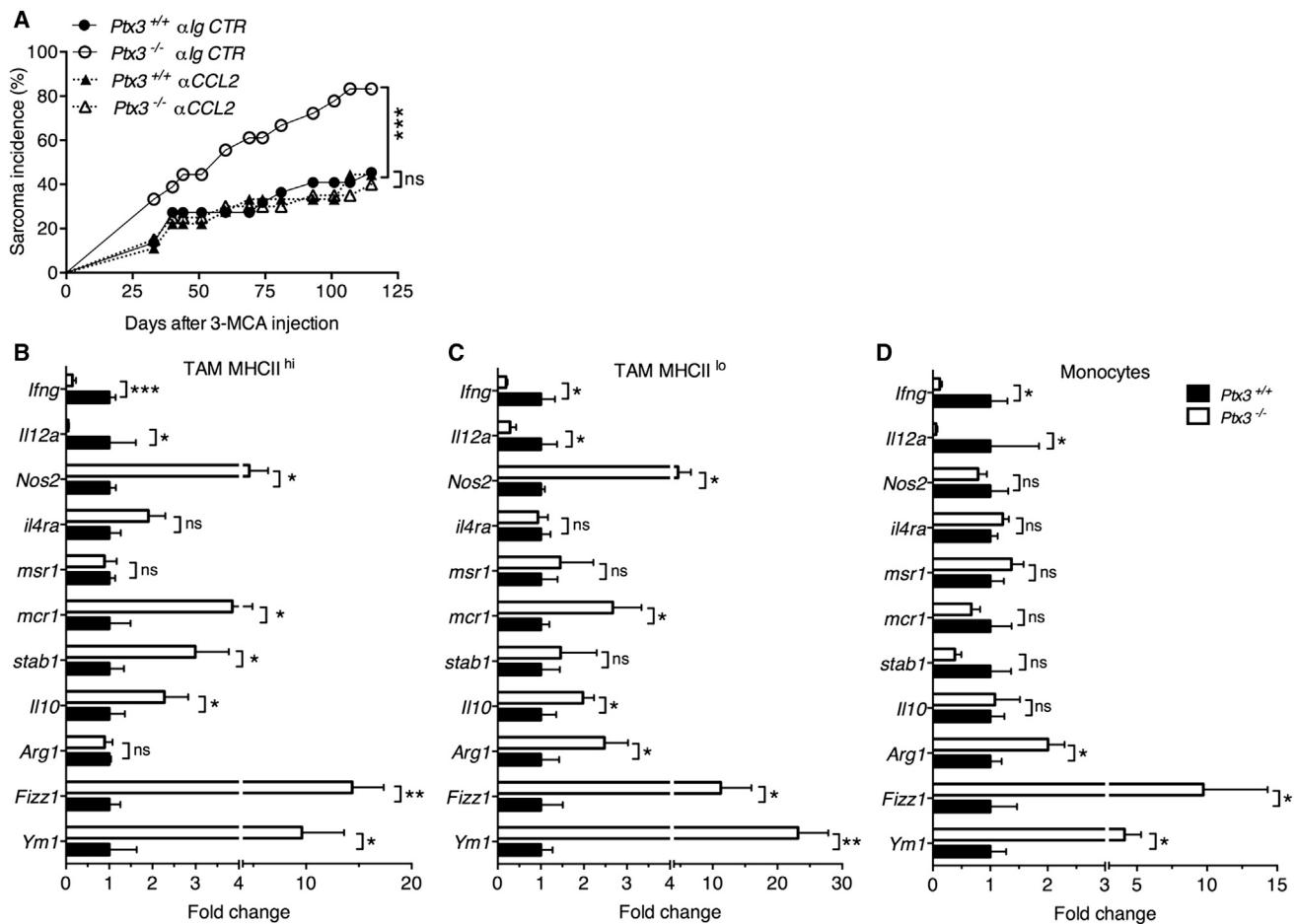
(E and F) Analysis by confocal microscopy of C5b-9 (E) and Factor H (F) in 3-MCA-tumors. Correlation analysis between Factor H and PTX3 expression was performed in wild-type tumors.

(G) Incidence of 3-MCA-sarcoma in *C3<sup>-/-</sup>* and/or *Ptx3<sup>-/-</sup>* and *Ptx3<sup>+/+</sup>* mice (n = 10–15).

(H) Analysis of the macrophage (F4/80<sup>+</sup>/CD11b<sup>+</sup>) infiltrate in 3-MCA-sarcoma by FACS in *C3<sup>-/-</sup>* and/or *Ptx3<sup>-/-</sup>* and *Ptx3<sup>+/+</sup>* mice.

(I) 3-MCA-sarcoma incidence in *Ptx3<sup>-/-</sup>* and *Ptx3<sup>+/+</sup>* mice treated with the C5aR antagonist PMX53 (n = 7).

(legend continued on next page)



**Figure 5. Role of CCL2 in Carcinogenesis and TAM Polarization in PTX3-Deficient Mice**

(A) Incidence of 3-MCA-sarcoma in *Ptx3*<sup>-/-</sup> and *Ptx3*<sup>+/+</sup> mice treated with anti-CCL2, or irrelevant mAb (n = 9–11) (100 µg/mouse three times/week) for the duration of the experiment. \*\*\*p < 0.001, paired Student's t test.

(B–D) Analysis by RT-PCR of selected M1 and M2 markers of macrophages (B and C) and monocytes (D) infiltrating 3-MCA-sarcoma in *Ptx3*<sup>-/-</sup> and *Ptx3*<sup>+/+</sup> mice (mean ± SEM). \*p < 0.05, \*\*p < 0.01, \*\*\*p < 0.001, unpaired Student's t test. In (B), (C), and (D), data were relative to 18S expression and normalized versus the mean of wild-type and expressed as mean ± SEM.

See also Table S2.

abolished the difference in 8-OH-dG and γ-H2Ax immunoreactivity between *Ptx3*<sup>-/-</sup> and *Ptx3*<sup>+/+</sup> tissues (Figures 6D and 6E). These results suggest that increased CRI in PTX3-deficient mice is associated with increased genetic instability as assessed by the frequency of *Trp53* mutations, oxidative DNA damage, and expression of DDR markers.

#### Methylation of the PTX3 Gene in Human Cancer

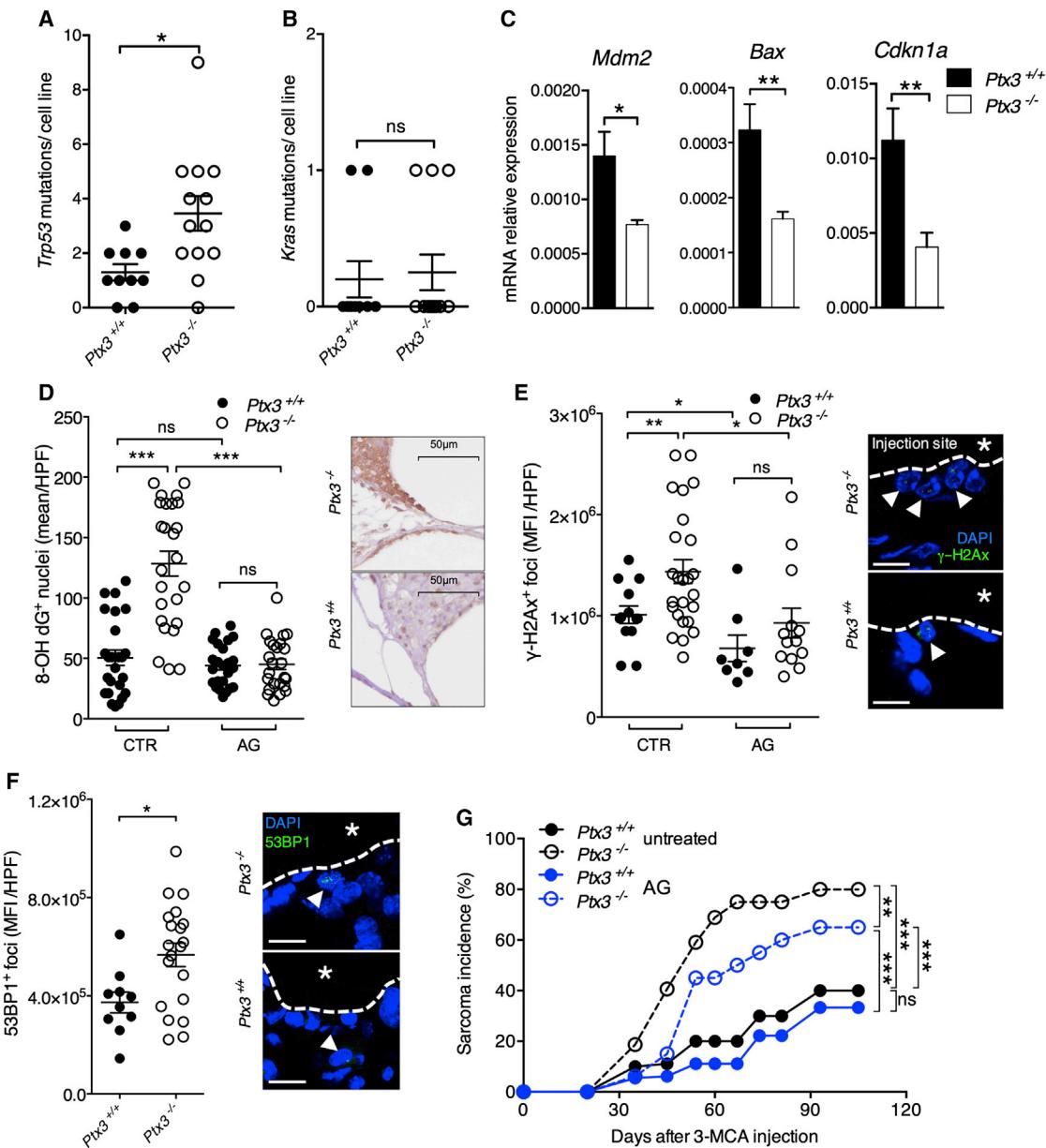
We next investigated whether results obtained in the mouse are indeed relevant to human cancer. Studies in human esophageal

squamous cell carcinoma had shown that the *PTX3* promoter is hypermethylated in this cancer and *PTX3* expression is inhibited (Wang et al., 2011). In addition, in silico bioinformatics analysis in the Epigenomics database (<http://www.ncbi.nlm.nih.gov/epigenomics>) showed *PTX3* gene methylation in colorectal cancer (CRC) (File: GSM801957). We thus analyzed the methylation status of the promoter and of a CpG island located in a putative enhancer encompassing the second exon (Akhtar-Zaidi et al., 2012) in a series of mesenchymal and epithelial cancers. As shown in Figure 7A, both the promoter and the CpG island are

(J) Immunofluorescence analysis of C3 deposition on *Ptx3*<sup>+/+</sup> and *Ptx3*<sup>-/-</sup> cells, in the presence of recombinant PTX3, SRC19-20, or their combination. Arrows indicate faint C3 immunostaining in *Ptx3*<sup>+/+</sup> cells in basal conditions or after addition of PTX3 in the assay.

(K) Quantification of C3 deposition evaluated by immunofluorescence. \*p < 0.05, \*\*p < 0.01, \*\*\*p < 0.001, unpaired Student's t test (A–F, H, and K), paired Student's t test (G and I). For (A), (B), (C), (E), (F), and (K), each dot represents the average of mean fluorescence intensity of ten regions of interest (ROI) per tumor (A–C, E, and F) or well (K) (mean ± SEM). Scale bars represent 150 µm.

See also Figure S4.



**Figure 6. PTX3 Deficiency Is Associated to Increased Genetic Instability**

(A and B) Number of *Trp53* (A) and *Kras* (B) mutations (mean  $\pm$  SEM) in cell lines isolated from *Ptx3*<sup>-/-</sup> and *Ptx3*<sup>+/+</sup> tumors. See also Table S3.

(C) Analysis of the expression of p53-regulated genes by RT-PCR in cell lines isolated from *Ptx3*<sup>-/-</sup> and *Ptx3*<sup>+/+</sup> tumors. Data were relative to GAPDH expression and represented as mean  $\pm$  SEM.

(D) Immunohistochemical analysis of 8-OH-dG 15 days after 3-MCA injection in *Ptx3*<sup>+/+</sup> and *Ptx3*<sup>-/-</sup> mice treated with AG. Left: data are reported as mean number  $\pm$  SEM of 8-OH-dG<sup>+</sup> nuclei per field. Right: representative immunofluorescence images are shown.

(E and F) Immunofluorescence analysis of  $\gamma$ -H2Ax (E) and 53BP1 (F) 15 days after 3-MCA injection in *Ptx3*<sup>+/+</sup> and *Ptx3*<sup>-/-</sup> mice treated with AG (E). Left panels: data are reported as MFI of  $\gamma$ -H2Ax (E) and 53BP1 (F) positive foci per field and expressed as (mean  $\pm$  SEM). Right panels: representative immunofluorescence images are shown.

(G) Sarcoma incidence in *Ptx3*<sup>+/+</sup> and *Ptx3*<sup>-/-</sup> mice treated with AG in the drinking water for the entire duration of the experiment. \*p < 0.05, \*\*p < 0.01, \*\*\*p < 0.001, unpaired (A-F) or paired (G) Student's t test.

highly methylated in leiomiosarcomas, desmoid tumors, CRC, and skin squamous cell carcinoma, in contrast with normal mesenchymal or epithelial tissues. We further investigated *PTX3* methylation and the functional consequences in CRC.

We analyzed *PTX3* methylation in five patients with high-grade adenomas and 40 patients with microsatellite stable CRC at different stage (I-IV). As shown in Figure 7B, the methylation of the promoter region and of the CpG island progressively

increased from normal colon epithelium, to adenomas and to CRC independently of stage. The healthy tissue from CRC patients was more methylated than normal colon epithelium. Similar results were obtained in colorectal tumor cell lines (Figures 7C and S5A), in which treatment with the methylation inhibitor 5-Aza-2'-deoxycytidine (5-AZA-dC) significantly reduced the methylation of both the promoter and the CpG island. In the presence of 5-AZA-dC, the expression of PTX3 mRNA was significantly increased in five different cell lines analyzed, both under basal conditions and upon stimulation with TNF $\alpha$  or IL-1 $\beta$  (Figures 7D and S5B). Rescue of PTX3 protein expression was confirmed in 3 different cell lines (Figures 7D and S5C). In contrast, treatment with 5-AZA-dC did not increase PTX3 expression in normal endothelial or normal colon epithelial cells stimulated with TNF $\alpha$  (Figures S5D and S5E). In addition, chromatin immunoprecipitation (ChIP) analysis showed that histone modifications associated to transcriptional activation (H3K4me3, H3K27ac, and H3K9ac) were strongly increased in the PTX3 promoter after treatment with 5-AZA-dC and TNF $\alpha$ , whereas trimethylation of H3K27 associated to gene repression was not modified (Figure S5F). Finally, the binding of NF- $\kappa$ B, c-Jun, and c-Fos to their consensus binding sites in the promoter and transcription factor binding sites where above the basal level only after cell treatment with 5-AZA-dC in association with a proinflammatory stimulus in CRC cells (Figure 7E). Collectively, these data suggest that the PTX3 gene is silenced by methylation in selected human tumors including CRC.

## DISCUSSION

The general objective of this investigation was to gain insight into the role of the humoral arm of innate immunity in CRI using the long pentraxin PTX3 as a paradigm (Bottazzi et al., 2010). We observed increased susceptibility to mesenchymal (3-MCA) and epithelial (DMBA/TPA) carcinogenesis in PTX3-deficient mice. The 3-MCA model was used for further analysis. This phenotype was associated with increased macrophage infiltration, cytokine production, angiogenesis, and genetic instability as revealed by an increased frequency of *Trp53* mutations, oxidative DNA damage, and expression of DDR markers in *Ptx3*<sup>-/-</sup> mice. Several lines of evidence suggest that PTX3 deficiency unleashes unrestrained Complement activation with production of C5a, CCL2 production, and enhanced recruitment of tumor-promoting macrophages. In selected human tumors (e.g., leiomyosarcomas and CRC) PTX3 expression was regulated epigenetically by methylation of the promoter region and of a putative enhancer. Thus, an essential component of the humoral arm of innate immunity and regulator of Complement activation acts as an extrinsic oncosuppressor by acting at the level of Complement-mediated, macrophage-sustained, tumor-promoting inflammation.

PTX3 is a humoral pattern recognition molecule essential for resistance against selected microorganisms including *A. fumigatus*, *P. aeruginosa* (Garlanda et al., 2002), *Klebsiella pneumonia* (Soares et al., 2006), and uropathogenic *Escherichia coli* (Jaillon et al., 2014). PTX3 is highly conserved in evolution and genetic evidence is consistent with a role of PTX3 in antimicrobial resistance in humans (Chiarini et al., 2010; Cunha et al.,

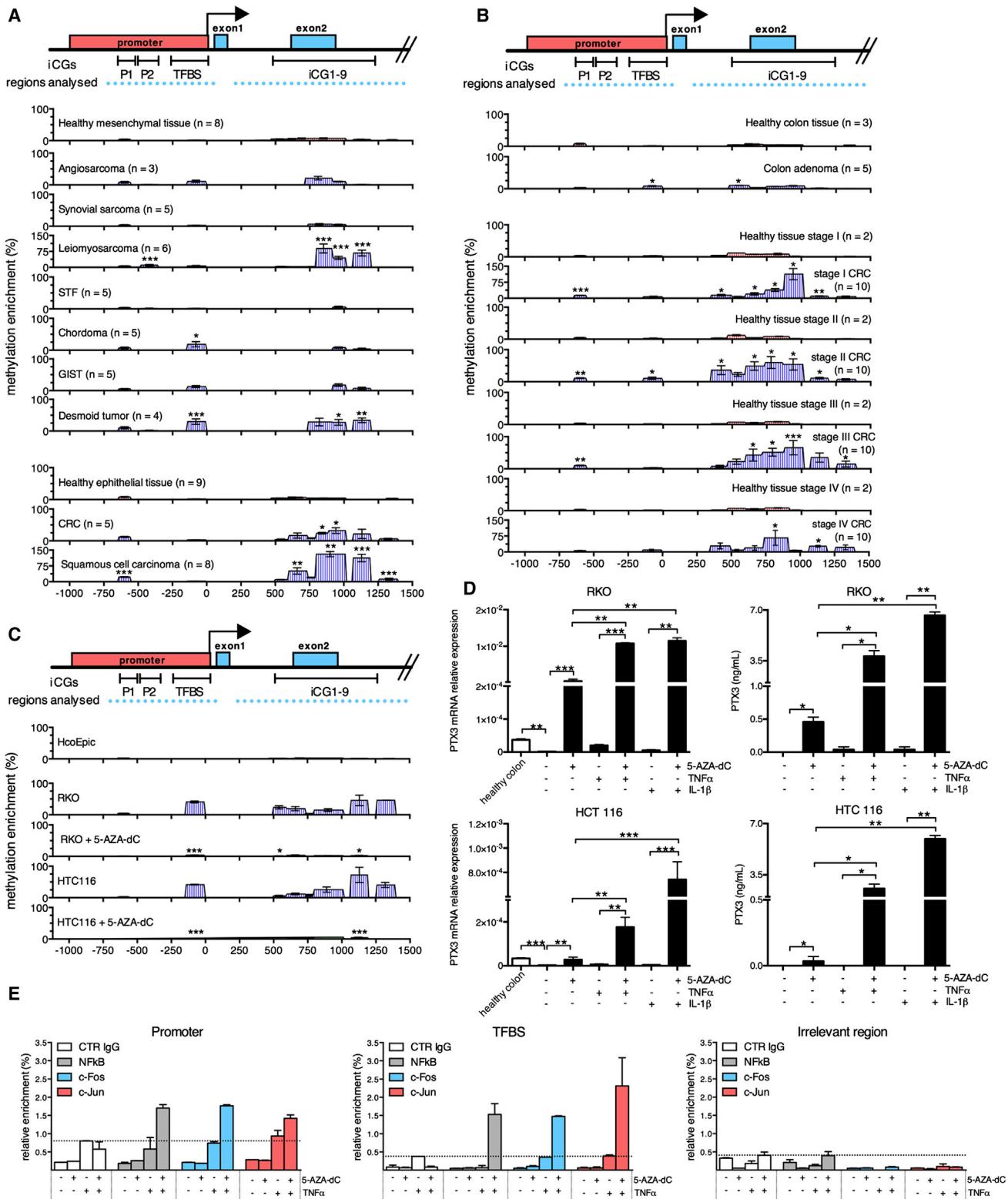
2014; Olesen et al., 2007). PTX3 has antibody-like properties, including recognition of microbial moieties (Jeannin et al., 2005), opsonization via Fc $\gamma$  receptors (Moalli et al., 2010), Complement activation and regulation (Deban et al., 2008; Inforzato et al., 2013), and regulation of inflammation (Deban et al., 2010; Lech et al., 2013; Norata et al., 2009; Salio et al., 2008; Soares et al., 2006). Given the fact that tumor-promoting inflammation is now recognized as an essential component of the tumor microenvironment (Coussens et al., 2013; Grivennikov et al., 2010; Hanahan and Weinberg, 2011; Mantovani et al., 2008), the increased susceptibility of PTX3-deficient mice to carcinogenesis is a reflection of its regulatory function on inflammation.

PTX3 binds P-selectin via its glycosidic component (Deban et al., 2010) and this accounts for the regulatory function on inflammation of PTX3 in selected P-selectin-dependent experimental models (Lech et al., 2013). It was therefore important to assess the relevance of the P-selectin pathway in the enhanced carcinogenesis observed in *Ptx3*<sup>-/-</sup> mice. P-selectin-deficient mice show no phenotype in terms of 3-MCA carcinogenesis. Moreover, PTX3/P-selectin-double-deficient mice showed enhanced 3-MCA carcinogenesis compared to controls. Thus, the P-selectin pathway is not involved in the enhanced susceptibility to carcinogenesis of PTX3-deficient mice.

PTX3 interacts with Complement at multiple levels (Inforzato et al., 2013). It binds C1q (Bottazzi et al., 1997) and activates or regulates (Nauta et al., 2003) the classic pathway of Complement activation. By interacting with ficolins and MBL it independently impinges upon the alternative pathway (Gout et al., 2011; Ma et al., 2009). PTX3 interacts and recruits Factor H by binding domains 19–20 and 7 through its N-terminal and glycosylated C-terminal domains, respectively, without interfering with Factor H capacity to negatively regulate the Complement cascade (Deban et al., 2008). We found an inverse relationship between C3 deposition and PTX3 presence in individual tumors. Conversely, the occurrence of Factor H was directly and significantly correlated with PTX3 deposition. In an effort to assess the actual significance of the Complement cascade in the onco-suppressive function of PTX3, double C3/PTX3-deficient mice were generated. Genetic inactivation of C3 strongly reduced, but did not completely ablate, the tumor-promoting function of PTX3 deficiency, suggesting that other mechanisms in addition to unleashed complement may contribute to increased CRI and susceptibility to cancer in PTX3-deficient mice.

A genetic approach to address the role of Factor H in vivo using double knockout mice is hindered by systemic C3 consumption and associated pathology in Factor H-deficient mice (Banda et al., 2013; Pickering et al., 2002). However, in an in vitro assay of C3 deposition on PTX3 competent and PTX3-incompetent cell lines, a recombinant Factor H inhibitor (SRC19-20) (Banda et al., 2013), which interferes with the binding between PTX3 and Factor H (Deban et al., 2008), increased C3 deposition on PTX3 producing cells and abolished the rescue of C3 deposition on *Ptx3*<sup>-/-</sup> cells obtained with recombinant PTX3. Thus, we infer that PTX3 deficiency unleashes Complement-dependent, C5a-mediated tumor-promoting inflammation, because of defective recruitment of the negative regulator Factor H.

Complement is not considered a canonical component of tumor-promoting inflammation (Coussens et al., 2013; Hanahan



**Figure 7. The PTX3 Gene Is Epigenetically Modified in Human Cancers**

(A) Analysis by methylated CpG island recovery assay (MIRA) of the percentage of methylation enrichment of the promoter region and of the PTX3 CpG island in human mesenchymal and epithelial tumors. CpG rich regions (iCGs) (P1, P2, transcription factor binding site [TFBS], and CpG island [iCG1-9]) and surrounding

(legend continued on next page)

and Weinberg, 2011; Mantovani et al., 2008). Complement can represent an effector mechanism against hematopoietic tumors activated by monoclonal antibodies (Golay et al., 2001). The generation of Complement C5a in a transplanted tumor enhanced growth by recruiting myeloid-derived suppressor cells and amplifying their T cell-directed suppressive abilities (Markiewski et al., 2008). However, no information based on genetic approaches concerning the role of Complement components in carcinogenesis is available. Here, we found that C3-deficient mice are strongly protected against 3-MCA (Figure 4G) and DMBA/TPA (Figure S4B) carcinogenesis, involving mesenchymal and epithelial cells, respectively. These results demonstrate that Complement is an essential component of tumor-promoting inflammation.

The Complement cascade can be activated by antibodies or by sensing tissue damage (Ricklin et al., 2010). Antibodies have been shown to orchestrate cancer-related inflammation in a model of epithelial carcinogenesis (HPV16 transgenic mice) by a remote control mechanism (Andreu et al., 2010). However, in that system the pathway driving tumor-promoting inflammation was Fc $\gamma$  receptor-dependent and Complement-independent (Andreu et al., 2010; de Visser et al., 2004). It is therefore tempting to speculate that an antibody-independent pathway (e.g., interaction with senescent or apoptotic cells) is responsible for Complement-mediated orchestration of tumor-promoting inflammation in the two models of mesenchymal (3-MCA) and epithelial (DMBA/TPA) carcinogenesis used in the present study. Interestingly, in the present study, C3-deficient mice showed reduced susceptibility to DMBA/TPA skin carcinogenesis but were not protected against HPV16-driven squamous carcinoma development (Andreu et al., 2010; de Visser et al., 2004). CRI can differ considerably in its drivers and components in tumors originating in different organs and tumors (Ruffell et al., 2012). It is likely that Complement may play different roles in different carcinogenesis contexts.

Macrophages are an essential component of tumor-promoting inflammation (Qian and Pollard, 2010; Sica and Mantovani, 2012). Tumors originating in a PTX3-deficient context were characterized by an increased frequency of TAM throughout the natural history of the tumor. Increased macrophage accumulation was completely abolished in C3-deficient or C3/PTX3-double-deficient hosts. PTX3 deficiency was associated with increased C5a and CCL2. Both molecules are potent monocyte attractants. Interestingly at the end point (tumor volume 3 cm<sup>3</sup>) macrophage, but not neutrophil, numbers were increased in PTX3-defi-

cient hosts, consistent with a function of CCL2 in sustaining TAM recruitment and tumor-promoting inflammation. Macrophages in tumors usually acquire a skewed M2-like phenotype oriented to tumor promotion (Sica and Mantovani, 2012). Chemoattractants, and CCL2 in particular, have been shown to skew macrophage function in an M2-like direction (Roca et al., 2009). In PTX3-deficient hosts, TAM showed enhanced expression of molecules associated with M2-like polarization, including Arg I, Fizz1, and Ym1. Skewing of macrophage function with, for instance, increased immunosuppressive activity and may contribute to enhanced carcinogenesis in PTX3-deficient hosts. Antibodies directed against CCL2 reverted the phenotype of PTX3-deficient mice in terms of enhanced susceptibility to 3-MCA carcinogenesis but did not protect PTX3-competent mice, as observed in other murine models and clinical trials (Löberg et al., 2007; Pienta et al., 2013) suggesting that the treatment is effective only under conditions of exacerbated CCL2 production and M2-polarization. Chemoattractants including C5a induce chemokines (Sozzani et al., 1995). Thus, uncontrolled Complement activation with C5a production is likely to be upstream of CCL2 that in turn recruits tumor-promoting TAM in increased numbers, with skewed M2-like phenotype, in PTX3-deficient hosts. However, the strong reduction of TAM in PTX3/C3-double-deficient mice was not sufficient to completely rescue the susceptibility to carcinogenesis of *Ptx3*<sup>-/-</sup> mice, suggesting the involvement of other cell types or mechanisms in addition to increased TAM in the observed phenotype.

Inflammatory mediators have been suggested to contribute to genetic events that cause cancer and to the genetic instability of tumors, although there is little formal evidence for that (Colotta et al., 2009). PTX3 is an extracellular molecule whose only cellular receptors identified so far are Fc $\gamma$  receptors (Lu et al., 2008; Moalli et al., 2010). PTX3 does not affect the growth and death of normal or neoplastic cells. PTX3-deficient mice showed increased susceptibility to carcinogenesis mediated by Complement-dependent tumor-promoting inflammation. Tumors developed in a PTX3-deficient context characterized by exacerbated inflammation had a higher frequency of *Trp53* mutations, increased DNA oxidative damage and higher expression of DDR markers (Figure 6). These findings indicate that exacerbated inflammation in the tumor microenvironment can indeed contribute to the genetic instability of cancer.

In an effort to assess the actual significance of the present findings to human cancer, supported by a bioinformatics analysis showing methylation of the PTX3 gene in CRC, we searched

regions analyzed by MIRA are shown below the structure of PTX3 gene. Healthy epithelial tissues were from the normal counterpart of colon cancer patients ( $n = 3$ ) and skin squamous carcinoma ( $n = 5$ ); healthy mesenchymal tissues were from leiomyosarcoma ( $n = 1$ ), angiosarcoma ( $n = 3$ ), chordoma ( $n = 3$ ), and desmoid tumor ( $n = 2$ ) patients.

(B) Analysis by MIRA of the percentage of methylation enrichment of the same regions in healthy colon epithelium, high grade adenomas, CRC, and healthy counterparts for each CRC stage.

(C) Analysis by MIRA of the percentage of methylation enrichment of the same regions in one normal colon epithelial cell line (HCoEpic) and two CRC cell lines (RKO and HCT116) treated or not with the inhibitor of DNA methylation 5-AZA-dC.

(D) PTX3 mRNA expression and protein production by RKO and HCT116 cell lines treated or not with 5-AZA-dC in the presence of TNF $\alpha$  20 ng/mL or IL-1 $\beta$  20 ng/mL. Healthy colon: primary human intestinal epithelial crypt cells isolated from surgical biopsies.

(E) Analysis by ChIP of NF- $\kappa$ B, c-Fos, and c-Jun binding to the promoter, TFBS, and an irrelevant region in RKO cell line treated or not with 5-AZA-dC in the presence of TNF $\alpha$  20 ng/ml. Data are expressed as percentage of enrichment relatively to input DNA normalized on a positive control and represented as mean  $\pm$  SEM. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , unpaired Student's t test.

See also Figure S5.

for epigenetic modifications of the *PTX3* gene in mesenchymal and epithelial tumors. In selected human tumors (e.g., leiomyosarcomas, CRC, and skin squamous cell carcinoma), we found that the *PTX3* gene was hypermethylated at the promoter region and in a CpG rich putative enhancer encompassing exon 2. The latter is indeed a functional enhancer (M.R., unpublished data). *PTX3* gene methylation was detected in CRC stages 1–4 as well as in adenomas, suggesting that this is an early event in the natural history of CRC (Vogelstein et al., 2013). Hypermethylation at the *PTX3* locus was also observed in CRC cell lines, where treatment with 5-AZA-dC rescued the *PTX3* production in response to inflammatory cytokines. ChIP analysis showed that only after treatment with 5-AZA-dC and an inflammatory stimulus of CRC cells, histone modifications associated to transcriptional activation occurred and the transcription factors responsible of *PTX3* expression (e.g., NF-κB, c-Jun, c-Fos) interacted with their binding sites in the *PTX3* promoter region. Thus the *PTX3* gene is silenced by hypermethylation in selected human tumors including CRC and this event occurs early in progression already at the level of adenomas.

The results presented here demonstrate that a regulatory component of the humoral arm of innate immunity, *PTX3*, acts as an extrinsic oncosuppressor gene in mouse and man. Although the connection between inflammation and cancer has long been perceived (Coussens et al., 2013; Grivennikov et al., 2010; Hanahan and Weinberg, 2011; Mantovani et al., 2008), we are not aware of a precedent for a humoral innate immunity effector molecule acting as a cancer gene. An additional general implication of the results reported here relates to Complement. The results obtained suggest that Complement is a key component of tumor-promoting cancer-related inflammation.

## EXPERIMENTAL PROCEDURES

### Cancer Models

Mice were injected s.c. with a single dose of 100 μg of 3-MCA (Sigma-Aldrich) dissolved in corn oil and assessed for tumor development over the course of 5 months. Skin carcinogenesis was performed as reported (Schoppa et al., 2011) in a two-step DMBA/TPA model. Procedures conformed to protocols approved by the Humanitas Clinical and Research Center in compliance with national and international law and policies. The study was approved by the Italian Ministry of Health.

### Analysis of Complement Deposition

Deposition of Complement components was analyzed by confocal microscopy on tumor tissue or cell lines. C3 deposition on tumor cell lines was performed by incubating cells with 10% normal mouse serum in gelatin-containing Veronal buffer, recombinant *PTX3*, and/or 33 nM Factor H inhibitor peptide SRC19-20 (Braunschweig and Józsi, 2011).

### Analysis of DNA Damage

Genomic DNA was extracted from 10 *Ptx3*<sup>+/+</sup> and 12 *Ptx3*<sup>-/-</sup> 3-MCA-induced sarcoma cell lines, and the direct sequencing of exons 5 to 8 of *Trp53* gene and of exons 1 and 2 of *Kras* gene was performed as previously reported (Watanabe et al., 1999). DNA damage marker (8-OH-dG) and DDR markers (γ-H2Ax and 53BP1) were analyzed by immunohistochemistry and immunofluorescence.

### Patients

Paraffin-embedded tissues from oncologic Caucasian patients, who underwent resection surgery at Humanitas Clinical and Research Center, were examined for *PTX3* gene epigenetic modification by methylated CpG island re-

cvery assay (MIRA). The Institutional Review Board of the Humanitas Clinical and Research Center approved this study (ICH-99/09).

Detailed procedures for in vivo studies, biochemical assays, FACS analysis, immunostaining and confocal analysis, quantitative PCR, statistical analysis, and all other methods are described in Extended Experimental Procedures.

## SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, five figures, and three tables and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2015.01.004>.

## AUTHOR CONTRIBUTIONS

E.B. designed and performed most experiments, analyzed the data, and drafted the manuscript. S.G. and M.R. performed experiments and analyzed data. R.P., P.K., E.Ba., and C.Gr. performed epigenetic studies and analyzed data. V.M., F.F., M.M., I.L., and M.R.G. performed experiments. J.D.L. provided reagents. A.D., S.T., F.P., and M.N. performed histological analysis and analyzed data. L.L., M.Ro., P.C., and G.B. provided human samples. S.J. contributed to the design of the experiments and data analysis. C.G. and A.M. conceived the study, directed research, designed experiments, analyzed data, and wrote the manuscript.

## ACKNOWLEDGMENTS

We thank N. Polentarutti for technical assistance, S. Meri for generously providing SRC19-20, and B. Bottazzi and P. Bianchi for discussion. The contribution of the European Commission (ERC to A.M., FP7-HEALTH-2011-ADITEC-N°280873), Ministero dell'Istruzione, dell'Università e della Ricerca (MIUR) (project FIRB RBLA039LSF), Associazione Italiana Ricerca sul Cancro (AIRC and AIRC 5x1000), and the Italian Ministry of Health is gratefully acknowledged.

Received: April 24, 2014

Revised: November 10, 2014

Accepted: December 19, 2014

Published: February 12, 2015

## REFERENCES

- Akhtar-Zaidi, B., Cowper-Sal-lari, R., Corradin, O., Saiakhova, A., Bartels, C.F., Balasubramanian, D., Myeroff, L., Lutterbaugh, J., Jarrar, A., Kalady, M.F., et al. (2012). Epigenomic enhancer profiling defines a signature of colon cancer. *Science* 336, 736–739.
- Andreu, P., Johansson, M., Affara, N.I., Pucci, F., Tan, T., Junankar, S., Korets, L., Lam, J., Tawfik, D., DeNardo, D.G., et al. (2010). FcRgamma activation regulates inflammation-associated squamous carcinogenesis. *Cancer Cell* 17, 121–134.
- Banda, N.K., Mehta, G., Ferreira, V.P., Cortes, C., Pickering, M.C., Pangburn, M.K., Arend, W.P., and Holers, V.M. (2013). Essential role of surface-bound complement factor H in controlling immune complex-induced arthritis. *J. Immunol.* 190, 3560–3569.
- Bottazzi, B., Vuoret-Craviari, V., Bastone, A., De Gioia, L., Matteucci, C., Peri, G., Spreafico, F., Pausa, M., D'Ettorre, C., Gianazza, E., et al. (1997). Multimer formation and ligand recognition by the long pentraxin PTX3. Similarities and differences with the short pentraxins C-reactive protein and serum amyloid P component. *J. Biol. Chem.* 272, 32817–32823.
- Bottazzi, B., Doni, A., Garlanda, C., and Mantovani, A. (2010). An integrated view of humoral innate immunity: pentraxins as a paradigm. *Annu. Rev. Immunol.* 28, 157–183.
- Braunschweig, A., and Józsi, M. (2011). Human pentraxin 3 binds to the complement regulator C4b-binding protein. *PLoS ONE* 6, e23991.
- Chiarini, M., Sabelli, C., Melotti, P., Garlanda, C., Savoldi, G., Mazza, C., Padoan, R., Plebani, A., Mantovani, A., Notarangelo, L.D., et al. (2010). PTX3

genetic variations affect the risk of *Pseudomonas aeruginosa* airway colonization in cystic fibrosis patients. *Genes Immun.* 11, 665–670.

Colotta, F., Allavena, P., Sica, A., Garlanda, C., and Mantovani, A. (2009). Cancer-related inflammation, the seventh hallmark of cancer: links to genetic instability. *Carcinogenesis* 30, 1073–1081.

Coussens, L.M., Zitvogel, L., and Palucka, A.K. (2013). Neutralizing tumor-promoting chronic inflammation: a magic bullet? *Science* 339, 286–291.

Cunha, C., Aversa, F., Lacerda, J.F., Busca, A., Kurzai, O., Grube, M., Löfller, J., Maertens, J.A., Bell, A.S., Inforzato, A., et al. (2014). Genetic PTX3 deficiency and aspergillosis in stem-cell transplantation. *N. Engl. J. Med.* 370, 421–432.

de Visser, K.E., Korets, L.V., and Coussens, L.M. (2004). Early neoplastic progression is complement independent. *Neoplasia* 6, 768–776.

Deban, L., Jarva, H., Lehtinen, M.J., Bottazzi, B., Bastone, A., Doni, A., Jokiranta, T.S., Mantovani, A., and Meri, S. (2008). Binding of the long pentraxin PTX3 to factor H: interacting domains and function in the regulation of complement activation. *J. Immunol.* 181, 8433–8440.

Deban, L., Russo, R.C., Sironi, M., Moalli, F., Scanziani, M., Zambelli, V., Cucovillo, I., Bastone, A., Gobbi, M., Valentino, S., et al. (2010). Regulation of leukocyte recruitment by the long pentraxin PTX3. *Nat. Immunol.* 11, 328–334.

Garlanda, C., Hirsch, E., Bozza, S., Salustri, A., De Acetis, M., Nota, R., Maccagno, A., Riva, F., Bottazzi, B., Peri, G., et al. (2002). Non-redundant role of the long pentraxin PTX3 in anti-fungal innate immune response. *Nature* 420, 182–186.

Garlanda, C., Bottazzi, B., Bastone, A., and Mantovani, A. (2005). Pentraxins at the crossroads between innate immunity, inflammation, matrix deposition, and female fertility. *Annu. Rev. Immunol.* 23, 337–366.

Golay, J., Lazzari, M., Facchinetto, V., Bernasconi, S., Borleri, G., Barbui, T., Rambaldi, A., and Introna, M. (2001). CD20 levels determine the in vitro susceptibility to rituximab and complement of B-cell chronic lymphocytic leukemia: further regulation by CD55 and CD59. *Blood* 98, 3383–3389.

Gout, E., Moriscot, C., Doni, A., Dumestre-Pérard, C., Lacroix, M., Pérard, J., Schoehn, G., Mantovani, A., Arlaud, G.J., and Thielens, N.M. (2011). M-ficolin interacts with the long pentraxin PTX3: a novel case of cross-talk between soluble pattern-recognition molecules. *J. Immunol.* 186, 5815–5822.

Grivennikov, S.I., Greten, F.R., and Karin, M. (2010). Immunity, inflammation, and cancer. *Cell* 140, 883–899.

Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. *Cell* 144, 646–674.

Inforzato, A., Doni, A., Barajon, I., Leone, R., Garlanda, C., Bottazzi, B., and Mantovani, A. (2013). PTX3 as a paradigm for the interaction of pentraxins with the complement system. *Semin. Immunol.* 25, 79–85.

Jaillon, S., Moalli, F., Ragnarsdóttir, B., Bonavita, E., Puthia, M., Riva, F., Bartoli, E., Nebuloni, M., Cvetko Krajinovic, L., Markotic, A., et al. (2014). The humoral pattern recognition molecule PTX3 is a key component of innate immunity against urinary tract infection. *Immunity* 40, 621–632.

Jeannin, P., Bottazzi, B., Sironi, M., Doni, A., Rusnati, M., Presta, M., Maina, V., Magistrelli, G., Haeuw, J.F., Hoeffel, G., et al. (2005). Complexity and complementarity of outer membrane protein A recognition by cellular and humoral innate immunity receptors. *Immunity* 22, 551–560.

Laoui, D., Van Overmeire, E., Di Conza, G., Aldeni, C., Keirsse, J., Morias, Y., Movahedi, K., Houbraken, I., Schouppe, E., Elkrim, Y., et al. (2014). Tumor hypoxia does not drive differentiation of tumor-associated macrophages but rather fine-tunes the M2-like macrophage population. *Cancer Res.* 74, 24–30.

Lech, M., Römmele, C., Gröbmayer, R., Eka Susanti, H., Kulkarni, O.P., Wang, S., Gröne, H.J., Uhl, B., Reichel, C., Krombach, F., et al. (2013). Endogenous and exogenous pentraxin-3 limits postischemic acute and chronic kidney injury. *Kidney Int.* 83, 647–661.

Loberg, R.D., Ying, C., Craig, M., Day, L.L., Sargent, E., Neeley, C., Wojno, K., Snyder, L.A., Yan, L., and Pienta, K.J. (2007). Targeting CCL2 with systemic delivery of neutralizing antibodies induces prostate cancer tumor regression in vivo. *Cancer Res.* 67, 9417–9424.

Lu, J., Marnell, L.L., Marjon, K.D., Mold, C., Du Clos, T.W., and Sun, P.D. (2008). Structural recognition and functional activation of Fc $\gamma$ R by innate pentraxins. *Nature* 456, 989–992.

Ma, Y.J., Doni, A., Hummelshøj, T., Honoré, C., Bastone, A., Mantovani, A., Thielens, N.M., and Garred, P. (2009). Synergy between ficolin-2 and pentraxin 3 boosts innate immune recognition and complement deposition. *J. Biol. Chem.* 284, 28263–28275.

Mantovani, A., Allavena, P., Sica, A., and Balkwill, F. (2008). Cancer-related inflammation. *Nature* 454, 436–444.

Markiewski, M.M., DeAngelis, R.A., Benencia, F., Ricklin-Lichtsteiner, S.K., Koutoulaki, A., Gerard, C., Coukos, G., and Lambris, J.D. (2008). Modulation of the antitumor immune response by complement. *Nat. Immunol.* 9, 1225–1235.

Moalli, F., Doni, A., Deban, L., Zelante, T., Zagarella, S., Bottazzi, B., Romani, L., Mantovani, A., and Garlanda, C. (2010). Role of complement and Fc $\gamma$  receptors in the protective activity of the long pentraxin PTX3 against *Aspergillus fumigatus*. *Blood* 116, 5170–5180.

Murray, P.J., Allen, J.E., Biswas, S.K., Fisher, E.A., Gilroy, D.W., Goerdt, S., Gordon, S., Hamilton, J.A., Ivashkiv, L.B., Lawrence, T., et al. (2014). Macrophage activation and polarization: nomenclature and experimental guidelines. *Immunity* 41, 14–20.

Nauta, A.J., Bottazzi, B., Mantovani, A., Salvatori, G., Kishore, U., Schwaeble, W.J., Gingras, A.R., Tzima, S., Vivanco, F., Egido, J., et al. (2003). Biochemical and functional characterization of the interaction between pentraxin 3 and C1q. *Eur. J. Immunol.* 33, 465–473.

Norata, G.D., Marchesi, P., Pulakazhi Venu, V.K., Pasqualini, F., Anselmo, A., Moalli, F., Pizzitola, I., Garlanda, C., Mantovani, A., and Catapano, A.L. (2009). Deficiency of the long pentraxin PTX3 promotes vascular inflammation and atherosclerosis. *Circulation* 120, 699–708.

Olesen, R., Wejse, C., Velez, D.R., Bisseye, C., Sodemann, M., Aaby, P., Rabna, P., Worwui, A., Chapman, H., Diatta, M., et al. (2007). DC-SIGN (CD209), pentraxin 3 and vitamin D receptor gene variants associate with pulmonary tuberculosis risk in West Africans. *Genes Immun.* 8, 456–467.

Pickering, M.C., Cook, H.T., Warren, J., Bygrave, A.E., Moss, J., Walport, M.J., and Botto, M. (2002). Uncontrolled C3 activation causes membranoproliferative glomerulonephritis in mice deficient in complement factor H. *Nat. Genet.* 31, 424–428.

Pienta, K.J., Machiels, J.P., Schrijvers, D., Alekseev, B., Shkolnik, M., Crabb, S.J., Li, S., Seetharam, S., Puchalski, T.A., Takimoto, C., et al. (2013). Phase 2 study of carlumab (CINTO 888), a human monoclonal antibody against CC-chemokine ligand 2 (CCL2), in metastatic castration-resistant prostate cancer. *Invest. New Drugs* 31, 760–768.

Qian, B.Z., and Pollard, J.W. (2010). Macrophage diversity enhances tumor progression and metastasis. *Cell* 141, 39–51.

Ricklin, D., Hajishengallis, G., Yang, K., and Lambris, J.D. (2010). Complement: a key system for immune surveillance and homeostasis. *Nat. Immunol.* 11, 785–797.

Roca, H., Varsos, Z.S., Sud, S., Craig, M.J., Ying, C., and Pienta, K.J. (2009). CCL2 and interleukin-6 promote survival of human CD11b+ peripheral blood mononuclear cells and induce M2-type macrophage polarization. *J. Biol. Chem.* 284, 34342–34354.

Ruffell, B., Affara, N.I., and Coussens, L.M. (2012). Differential macrophage programming in the tumor microenvironment. *Trends Immunol.* 33, 119–126.

Salio, M., Chimenti, S., De Angelis, N., Molla, F., Maina, V., Nebuloni, M., Pasqualini, F., Latini, R., Garlanda, C., and Mantovani, A. (2008). Cardioprotective function of the long pentraxin PTX3 in acute myocardial infarction. *Circulation* 117, 1055–1064.

Schioppa, T., Moore, R., Thompson, R.G., Rosser, E.C., Kulbe, H., Nedospasov, S., Mauri, C., Coussens, L.M., and Balkwill, F.R. (2011). B regulatory cells and the tumor-promoting actions of TNF- $\alpha$  during squamous carcinogenesis. *Proc. Natl. Acad. Sci. USA* 108, 10662–10667.

Sica, A., and Mantovani, A. (2012). Macrophage plasticity and polarization: in vivo veritas. *J. Clin. Invest.* 122, 787–795.

Sierra-Filardi, E., Nieto, C., Domínguez-Soto, A., Barroso, R., Sánchez-Mateos, P., Puig-Kroger, A., López-Bravo, M., Joven, J., Ardavín, C., Rodríguez-Fernández, J.L., et al. (2014). CCL2 shapes macrophage polarization by GM-CSF and M-CSF: identification of CCL2/CCR2-dependent gene expression profile. *J. Immunol.* 192, 3858–3867.

Soares, A.C., Souza, D.G., Pinho, V., Vieira, A.T., Nicoli, J.R., Cunha, F.Q., Mantovani, A., Reis, L.F., Dias, A.A., and Teixeira, M.M. (2006). Dual function of the long pentraxin PTX3 in resistance against pulmonary infection with *Klebsiella pneumoniae* in transgenic mice. *Microbes Infect.* 8, 1321–1329.

Sozzani, S., Sallusto, F., Luini, W., Zhou, D., Piemonti, L., Allavena, P., Van Damme, J., Valitutti, S., Lanzavecchia, A., and Mantovani, A. (1995). Migration of dendritic cells in response to formyl peptides, C5a, and a distinct set of chemokines. *J. Immunol.* 155, 3292–3295.

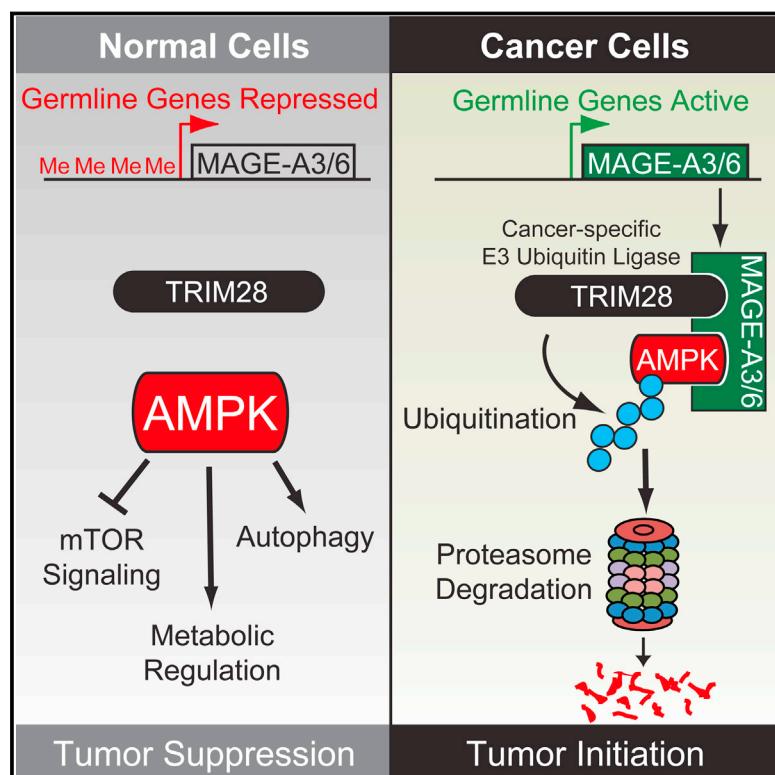
Vogelstein, B., Papadopoulos, N., Velculescu, V.E., Zhou, S., Diaz, L.A., Jr., and Kinzler, K.W. (2013). Cancer genome landscapes. *Science* 339, 1546–1558.

Wang, J.X., He, Y.L., Zhu, S.T., Yang, S., and Zhang, S.T. (2011). Aberrant methylation of the 3q25 tumor suppressor gene PTX3 in human esophageal squamous cell carcinoma. *World J. Gastroenterol.* 17, 4225–4230.

Watanabe, H., Shimokado, K., Asahara, T., Dohi, K., and Niwa, O. (1999). Analysis of the c-myc, K-ras and p53 genes in methylcholanthrene-induced mouse sarcomas. *Jpn. J. Cancer Res.* 90, 40–47.

# Degradation of AMPK by a Cancer-Specific Ubiquitin Ligase

## Graphical Abstract



## Authors

Carlos T. Pineda,  
Saumya Ramanathan, ...  
Michael A. White, Patrick Ryan Potts

## Correspondence

[ryan.potts@utsouthwestern.edu](mailto:ryan.potts@utsouthwestern.edu)

## In Brief

A cancer-specific E3 ubiquitin ligase ubiquitinates and degrades AMPK, resulting in downregulation of autophagy and increased mTOR signaling. This regulatory axis demonstrates how altered cellular metabolism can act as an oncogenic driver in cancer.

## Highlights

- MAGE-A3/6 are normally testis restricted and aberrantly expressed in many cancers
- MAGE-A3/6 are driver oncogenes competent to transform cells
- MAGE-A3/6-TRIM28 ubiquitinates and degrades the AMPK tumor suppressor
- MAGE-A3/6-TRIM28 suppresses autophagy and potentiates mTOR signaling

# Degradation of AMPK by a Cancer-Specific Ubiquitin Ligase

Carlos T. Pineda,<sup>1,4</sup> Saumya Ramanathan,<sup>1,4</sup> Klementina Fon Tacer,<sup>1</sup> Jenny L. Weon,<sup>1</sup> Malia B. Potts,<sup>2</sup> Yi-Hung Ou,<sup>2</sup> Michael A. White,<sup>2</sup> and Patrick Ryan Potts<sup>1,3,\*</sup>

<sup>1</sup>Department of Physiology, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA

<sup>2</sup>Department of Cell Biology, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA

<sup>3</sup>Departments of Pharmacology and Biochemistry, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA

<sup>4</sup>Co-first author

\*Correspondence: ryan.potts@utsouthwestern.edu

<http://dx.doi.org/10.1016/j.cell.2015.01.034>

## SUMMARY

**AMP-activated protein kinase (AMPK) is a master sensor and regulator of cellular energy status. Upon metabolic stress, AMPK suppresses anabolic and promotes catabolic processes to regain energy homeostasis. Cancer cells can occasionally suppress the growth-restrictive AMPK pathway by mutation of an upstream regulatory kinase. Here, we describe a widespread mechanism to suppress AMPK through its ubiquitination and degradation by the cancer-specific MAGE-A3/6-TRIM28 ubiquitin ligase. MAGE-A3 and MAGE-A6 are highly similar proteins normally expressed only in the male germline but frequently re-activated in human cancers. MAGE-A3/6 are necessary for cancer cell viability and are sufficient to drive tumorigenic properties of non-cancerous cells. Screening for targets of MAGE-A3/6-TRIM28 revealed that it ubiquitinates and degrades AMPK $\alpha$ . This leads to inhibition of autophagy, activation of mTOR signaling, and hypersensitization to AMPK agonists, such as metformin. These findings elucidate a germline mechanism commonly hijacked in cancer to suppress AMPK.**

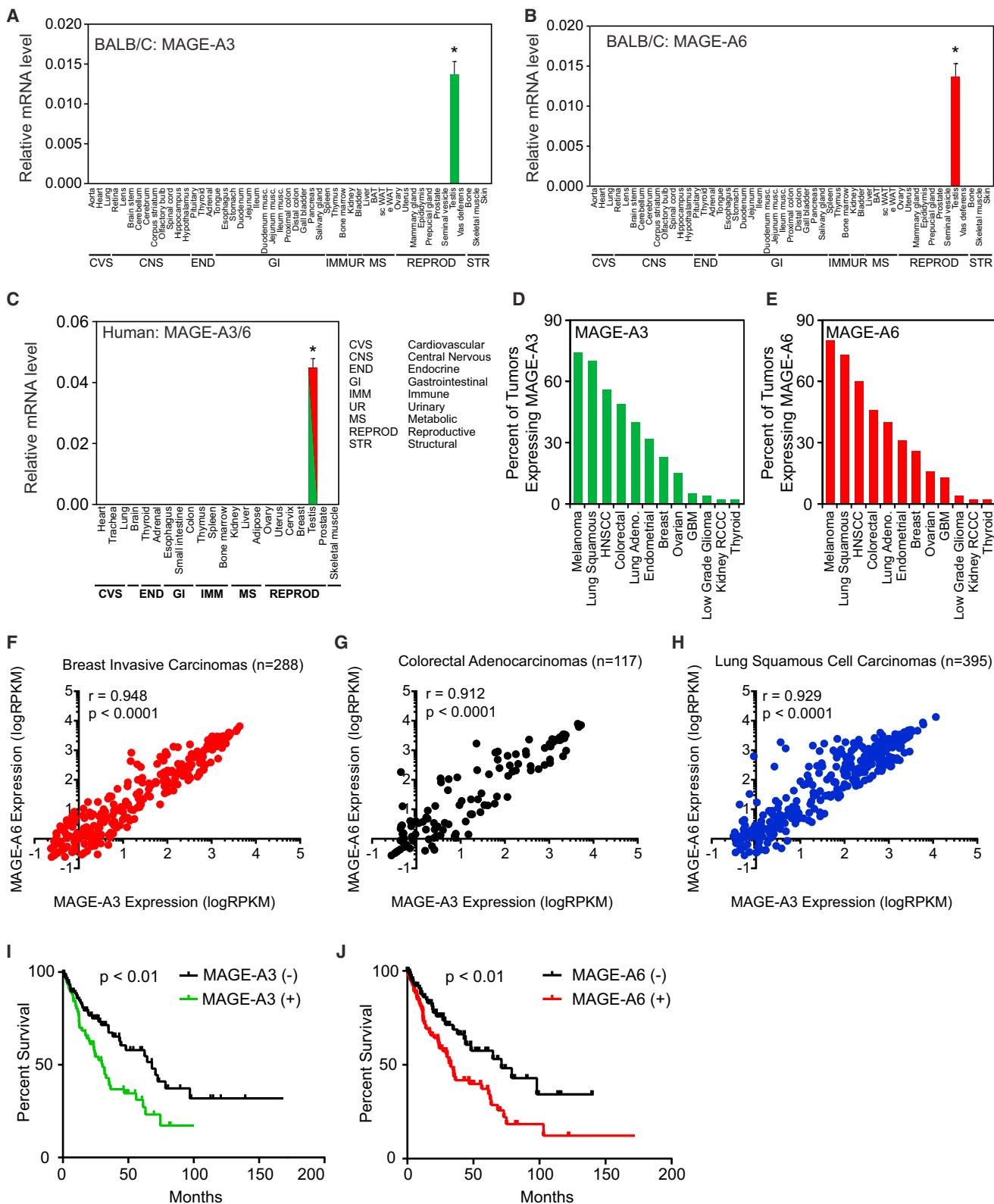
## INTRODUCTION

Cells must coordinate multiple metabolic processes in order to balance their energy usage with nutrient availability. One prominent way that this balance is accomplished is through the activity of the AMP-activated protein kinase (AMPK). AMPK is a heterotrimeric kinase comprised of catalytic  $\alpha$  and regulatory  $\beta$  and  $\gamma$  subunits that is regulated by the cellular concentrations of ATP, ADP, and AMP (Hardie et al., 2012b). When cellular levels of ATP fall and ADP/AMP rise, ATP that is bound to the  $\gamma$  subunit is replaced by ADP and/or AMP, resulting in activation of the catalytic kinase subunit (Landgraf et al., 2013; Suter et al., 2006). Once activated, AMPK generally opposes anabolic energy-consuming pathways while promoting catabolic ATP-generating pathways. For example, AMPK inhibits ACC1 and mTOR to block fatty acid and protein synthesis, respectively, while at

the same time it promotes autophagy via multiple pathways involving mTOR, ULK1, and VPS34 (Egan et al., 2011; Gwinn et al., 2008; Hardie et al., 2012b; Kim et al., 2011, 2013). In addition to changes in energy levels, upstream kinases such as LKB1/STK11 and CaMKK regulate AMPK activity by phosphorylation of its activation loop at T172 (Hawley et al., 2003, 2005; Shaw et al., 2004; Woods et al., 2005).

Although AMPK may in some cases promote late-stage tumor growth (Laderoute et al., 2014), multiple lines of evidence suggest AMPK has critical tumor suppressor activities in both humans and experimental models, including mice (Hardie and Alessi, 2013; Shackelford and Shaw, 2009). For example, knockout of AMPK $\alpha$ 1 in the mouse accelerates development of c-Myc-driven lymphomas (Faubert et al., 2013). AMPK's role in suppressing tumor initiation and progression is multifaceted, including growth suppression by inhibiting synthesis of cellular macromolecules (Hardie et al., 2012b), particularly through downregulating the mTOR signaling pathway (Gwinn et al., 2008; Inoki et al., 2003), and promoting cell-cycle arrest through stabilizing p53 and cyclin-dependent kinase inhibitors (Imamura et al., 2001; Jones et al., 2005; Liang et al., 2007). Additionally, AMPK can oppose the Warburg effect in favor of oxidative phosphorylation through upregulating oxidative enzymes and promoting mitochondrial biogenesis (Cantó et al., 2009; Winder et al., 2000). Furthermore, AMPK has recently been shown to inhibit epithelial-to-mesenchymal transition (EMT) by modulating the Akt-MDM2-Foxo3 signaling axis (Chou et al., 2014).

Given the importance of metabolic control and AMPK's role as master sensor and regulator of cellular energy, it is not surprising that this signaling axis is de-regulated in a variety of disease states, including cancer (Hardie and Alessi, 2013; Shackelford et al., 2009). For example, in approximately 20% of lung adenocarcinomas and cervical cancers, signaling through this axis is reduced by loss-of-function mutation or deletion of Lkb1/Stk11 (Matsumoto et al., 2007; Sanchez-Cespedes et al., 2002; Wingo et al., 2009). Additionally, AMPK levels have been shown to be reduced in some cases of hepatocellular carcinomas and BRAF V600E can downregulate AMPK signaling through inhibition of Lkb1/Stk11 in melanomas (Esteve-Puig et al., 2009; Lee et al., 2012; Zheng et al., 2009; Zheng et al., 2013). From these multiple lines of converging evidence on AMPK's critical role in tumor suppression, there is great interest in the utilization of compounds that stimulate AMPK activity, such as metformin, in the



(legend on next page)

prevention and treatment of cancer and many clinical trials are ongoing (Hadad et al., 2011; Hardie et al., 2012a; Niraula et al., 2012; Pernicova and Korbonits, 2014).

Melanoma antigen (MAGE) genes are conserved in all eukaryotes, encode for proteins with a common MAGE homology domain, and have rapidly expanded to comprise almost 50 unique genes in humans (Chomez et al., 2001; Feng et al., 2011). Approximately two-thirds of human MAGEs are considered cancer-testis antigens because they are normally restricted to expression in the testis but are aberrantly re-expressed in cancer and have antigenic properties (Simpson et al., 2005). The functional significance of MAGEs in tumors is not well understood, but accumulating evidence supports their importance. For example, knockdown of MAGE-A3/6 impairs tumor growth in mice, whereas expression of MAGE-A3 in MAGE-negative cells drives tumor growth and metastasis in vivo (Liu et al., 2008; Yang et al., 2007). Importantly, we recently showed that a defining characteristic of MAGE proteins is their ability to bind and potentiate the activity of specific E3 ubiquitin ligases (Doyle et al., 2010). For example, MAGE-L2 binds to the TRIM27 ubiquitin ligase and promotes ubiquitination of the WASH actin assembly complex to facilitate endosomal protein recycling (Hao et al., 2013).

Here, we present evidence for a regulatory axis engaged in cancer cells that downregulates AMPK through ubiquitination and degradation of AMPK $\alpha$ 1 by the normally testis-restricted MAGE-A3/6-TRIM28 E3 ubiquitin ligase complex activated in cancer. These findings identify a widespread mechanism for downregulating AMPK signaling during tumorigenesis and elucidate an unanticipated mechanism of action for oncogenic MAGE cancer-testis antigens.

## RESULTS

### MAGE-A3 and MAGE-A6 Are Physiologically Restricted to Expression in the Testis but Are Aberrantly Expressed in Cancer

MAGE-A3 and MAGE-A6 are highly similar, neighboring genes located on the X chromosome that encode proteins with 96% identity (Figures S1A and S1B). Given their homology and functional redundancy (see below), we refer to these genes as simply MAGE-A3/6 herein. To thoroughly examine the expression pattern of MAGE-A3/6, we analyzed their expression by quantitative RT-PCR (qRT-PCR) in >50 mouse tissues from two strains of mice (C57BL/6 and BALB/C). Consistent with previous findings (De Plaen et al., 1994), mouse MAGE-A3/6 were completely restricted to expression in the testis with no detectable expression in any other tissue (Figures 1A, 1B, S1C, and S1D). We extended these analyses to a panel of >20 human tissues and

found that human MAGE-A3/6 are similarly restricted to expression only in the human testis (Figure 1C). Like other cancer-testis antigen genes, MAGE-A3/6 have been reported to be aberrantly expressed in tumors (Jang et al., 2001; Shantha Kumara et al., 2012). Our analysis of a variety of different tumor types from patients revealed that MAGE-A3/6 are commonly expressed in many cancer types, including breast invasive carcinomas (25%), colon adenocarcinomas (50%), and lung squamous cell carcinomas (75%; Figures 1D and 1E). Additionally, expression of MAGE-A3 and MAGE-A6 was significantly correlated in breast invasive carcinomas, colon adenocarcinomas, and lung squamous cell carcinomas (Figures 1F–1H). However, expression of MAGE-A3 was not significantly correlated with expression of unrelated MAGE-A11 or MAGE-B2 (Figures S1E and S1F), suggesting that MAGE-A3 and MAGE-A6 expression is selectively coordinated. Furthermore, to determine whether MAGE-A3/6 expression correlates with patient outcome, we analyzed whether expression of MAGE-A3/6 correlates with overall survival. Indeed, patients with lung squamous cell carcinomas expressing MAGE-A3 or MAGE-A6 have a significant decrease in overall survival time (Figures 1I and 1J). Patients with tumors expressing MAGE-A3 had a >50% reduced overall survival time compared to patients with MAGE-A3-negative tumors (30 versus 69 months, respectively; 2.0 hazard ratio; Table S1). Similarly, patients with tumors expressing MAGE-A6 had a >50% reduced survival time (33 versus 71 months; hazard ratio of 1.9; Table S1). Together, these results suggest MAGE-A3/6 are physiologically restricted to the testis in both humans and mice but are frequently found in a wide variety of cancer types, and their expression correlates with poor patient prognosis.

### MAGE-A3/6 Are Required for Cancer Cell Viability and Function as Oncogenes

MAGE-A3/6 could be “passenger” genes that have little functional role or significance in tumorigenesis and are simply biomarkers. Alternatively, MAGE-A3/6 may be oncogenic “driver” genes that are involved in promoting tumor initiation and/or progression. To determine if MAGE-A3/6 have important functional roles in cancer cells, we examined whether patient-derived breast, colon, and lung cancer cells require the expression of MAGE-A3/6 for viability. Indeed, knockdown of MAGE-A3/6 using two independent small interfering RNAs (siRNAs) (Figures S2A and S2B) in multiple lung (HCC193, H1648, and H2126), breast (HCC1806 and SK-BR-3), and colon (HCT116 and HT29) cancer cells resulted in a significant decrease in cell viability and clonogenic survival (Figures 2A–2C, S2C, and S2D). Importantly, these effects are likely on target, because MAGE-A3/6 siRNAs do not significantly alter the viability of MAGE-A3/6-negative cells (HCC1143 and DLD1), even though

**Figure 1. MAGE-A3 and MAGE-A6 Are Normally Restricted to Expression in the Testis but Are Aberrantly Expressed in Cancer and Predict Poor Patient Prognosis**

(A and B) qRT-PCR analysis ( $n = 3$ ) of the normalized expression of mouse MAGE-A3 (A) and MAGE-A6 (B) in the indicated tissues from BALB/C mice.

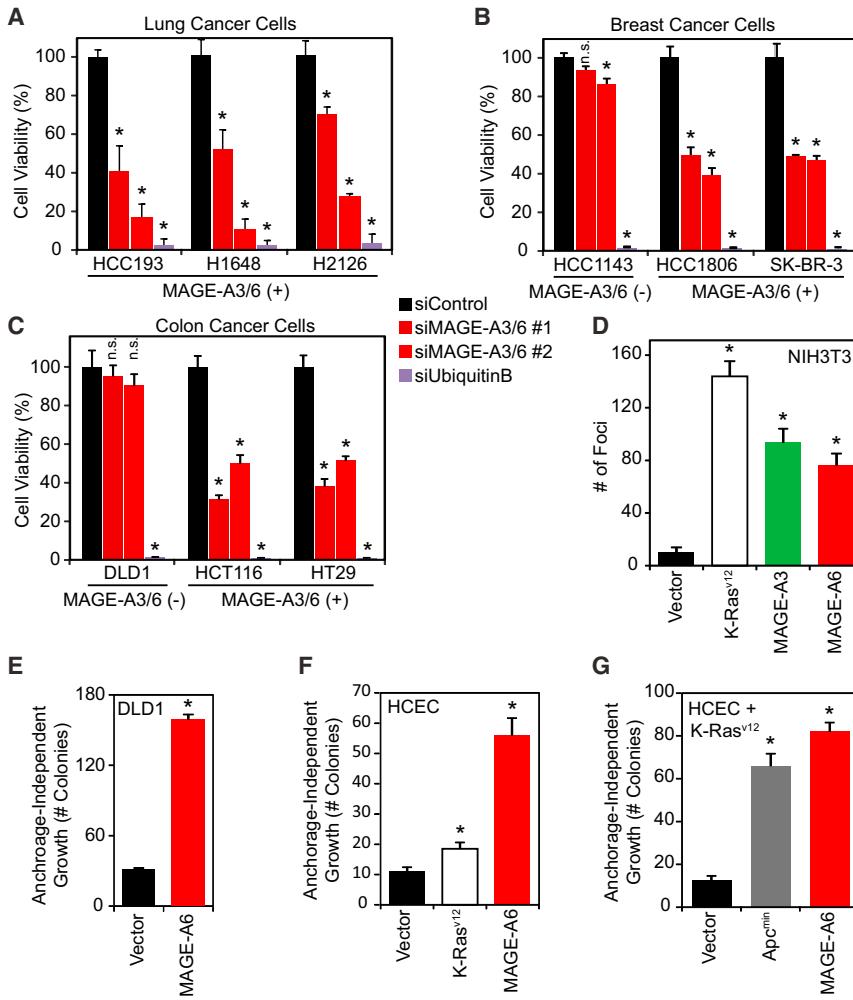
(C) qRT-PCR analysis ( $n = 3$ ) of the normalized expression of human MAGE-A3/6 (one primer set detects both) in the indicated human tissues.

(D and E) Percentage of patient tumors expressing MAGE-A3 (D) and MAGE-A6 (E) is shown.

(F–H) MAGE-A3 and MAGE-A6 are co-expressed in breast invasive carcinomas (F), colorectal adenocarcinomas (G), and lung squamous cell carcinomas (H).

(I and J) Expression of MAGE-A3 (C) or MAGE-A6 (D) in patients with lung squamous carcinomas correlates with poor overall survival.

Data are represented as the mean  $\pm$  SD. Asterisks indicate  $p < 0.05$ . See also Figure S1 and Table S1.



**Figure 2. MAGE-A3/6 Are Potent Oncogenes Whose Expression Is Necessary for Viability of Cancer Cells and Is Sufficient to Transform Cells**

(A–C) MAGE-A3/6 depletion reduces viability of MAGE-A3/6-expressing tumor cell lines. Lung (A), breast (B), and colon (C) cancer cells were treated with siControl, two distinct siMAGE-A3/6, or cytotoxic siUbiquitinB for a transfection control. Cell viability was measured by MTT assay.

(D) MAGE-A3 and MAGE-A6 have oncogenic activity. NIH 3T3 fibroblasts were transfected with MAGE-A3, MAGE-A6, or mutant K-Ras<sup>v12</sup> as a positive control, and foci formation was assayed.

(E) MAGE-A6 promotes anchorage-independent growth of DLD1 colon cancer cells. MAGE-A-negative DLD1 cells stably expressing vector or MAGE-A6 were assayed for anchorage-independent growth in soft agar colony formation assays.

(F and G) MAGE-A6 promotes anchorage-independent growth of non-transformed, immortalized human colonic epithelial cells (HCECs) without (F) or with (G) expression of mutant K-Ras<sup>v12</sup>. The indicated HCEC cells were assayed for anchorage-independent growth in soft agar colony formation assays.

Data ( $n = 3$ ) are represented as the mean  $\pm$  SD. Asterisks indicate  $p < 0.05$ . See also Figure S2.

the cytotoxic siRNA targeting ubiquitin B was just as lethal in all cell lines (Figures 2A–2C). These results suggest that upon expression of MAGE-A3/6, cells become dependent on their expression for viability, similar to other reports of “oncogene addiction” (Weinstein, 2002).

To determine whether MAGE-A3/6 are indeed oncogenic driver genes, we analyzed their activity in several classical assays. First, expression of either MAGE-A3 or MAGE-A6 significantly stimulated foci formation of NIH 3T3 cells (Figure 2D). Furthermore, MAGE-A6 promoted other hallmarks of cancer, such as anchorage-independent growth of the MAGE-A3/6-negative DLD1 colon cancer cells (Figure 2E). Finally, to more stringently assay the oncogenic activity of MAGE-A3/6, we determined the ability of MAGE-A6 to promote tumorigenic phenotypes in non-transformed, human colonic epithelial cells (HCECs) derived from normal colon biopsies and immortalized with CDK4 and hTERT (Roig et al., 2010). Remarkably, stable expression of MAGE-A6 alone was sufficient to drive transformation of these cells, resulting in robust anchorage-independent growth in soft agar (Figure 2F). In this setting, the oncogenic activity of MAGE-A6 was even more robust than expression of the bona fide K-Ras<sup>v12</sup> oncogene (Figure 2F). Additionally, even in

the context of mutant K-Ras<sup>v12</sup> expression, MAGE-A6 was still competent to drive anchorage-independent growth of HCEC cells to a similar degree as the *Apc<sup>min</sup>* oncogene (Figure 2G). Consistent with these findings, MAGE-A3/6 drives increased tumor growth and metastasis in an orthotopic xenograft mouse model

of thyroid cancer (Liu et al., 2008). Collectively, these findings suggest that MAGE-A3/6 are potent driver oncogenes that have therapeutic potential.

### The MAGE-A3/6-TRIM28 E3 Ubiquitin Ligase Ubiquitinates and Degrades AMPK $\alpha$ 1

Previously, we reported that MAGE proteins, including MAGE-A3/6, form complexes with specific E3 ubiquitin ligases to regulate ubiquitination (Doyle et al., 2010; Hao et al., 2013). MAGE-A3/6 specifically bind to the TRIM28 E3 ubiquitin ligase, also known as KAP1 (Doyle et al., 2010). We and others have previously shown that MAGE-A proteins can inhibit the critical p53 tumor suppressor, including via MAGE-A-TRIM28-induced ubiquitination and proteasome-mediated degradation of p53 (Doyle et al., 2010; Marcar et al., 2010; Monte et al., 2006; Wang et al., 2005; Yang et al., 2007). However, several of the cell lines dependent on MAGE-A3/6 for viability (Figures 2A–2C) are p53 null (HCC1806 and H1648) or mutant (HCC193, SK-BR-3, and HT-29). Additionally, MAGE-A3/6 stimulated anchorage-independent growth in p53 mutant DLD1 colon cancer cells (Figure 2E). Furthermore, expression of MAGE-A3/6 did not inversely correlate with p53 mutational status ( $\chi^2 = 0.17$ ;

$p = 0.98$ ). Therefore, the MAGE-A3/6-TRIM28 E3 ubiquitin ligase may have additional targets relevant to its function in cancer cells.

To identify additional targets of the MAGE-TRIM28 E3 ubiquitin ligase, we screened for its direct substrates by performing *in vitro* ubiquitination reactions on protein microarrays containing >9,000 SF9-purified, recombinant proteins (Figure 3A). This screen yielded 19 potential MAGE-TRIM28 substrates (Table S2) whose ubiquitination were significantly ( $p < 0.05$ ) increased by MAGE-TRIM28. To validate the results of the screen, candidates were tested for regulation by MAGE-A3/6-TRIM28, with AMPK $\alpha$ 1 being the most consistent and robust target of MAGE-A3/6-TRIM28 (Figures 3B–3D). Multiple siRNAs targeting MAGE-A3/6 or TRIM28 decreased ubiquitination of AMPK $\alpha$ 1 (Figures 3B and S2F). Expression of MAGE-A3 in MAGE-A3/6-negative cells promoted ubiquitination of AMPK $\alpha$ 1 (Figure 3C). Furthermore, knockdown of MAGE-A3/6 or TRIM28 resulted in increased AMPK $\alpha$ 1 protein levels (Figure 3D), without affecting mRNA levels (Figure 3E). Additional subunits of the AMPK holoenzyme complex, such as AMPK $\beta$ 1 and AMPK $\gamma$ 1, were correspondingly elevated (Figure S2G). In contrast, expression of MAGE-A3 in MAGE-A3/6-negative cells decreased AMPK $\alpha$ 1 protein levels (Figure 3F), without affecting mRNA levels (Figure S2E). Notably, AMPK $\alpha$ 1 protein levels could be rescued by addition of the MG132 proteasome inhibitor (Figure 3F), suggesting that MAGE-A3/6-TRIM28 ubiquitination of AMPK $\alpha$ 1 leads to its degradation by the proteasome.

To determine if MAGE-A3/6 enhances degradation of AMPK $\alpha$ 1 by TRIM28, as is the case for p53 (Doyle et al., 2010), or if MAGE-A3/6 is required for specifying AMPK $\alpha$ 1 degradation by TRIM28, we examined whether TRIM28 regulates AMPK $\alpha$ 1 levels in MAGE-A3/6-negative cells. Knockdown of TRIM28 in MAGE-A3/6-negative cells had no effect on AMPK $\alpha$ 1 levels (Figure 3G), suggesting that MAGE-A3/6 is required for AMPK $\alpha$ 1 degradation by TRIM28. Consistent with this notion, the AMPK $\alpha$ 1 $\beta$ 1 $\gamma$ 1 holoenzyme complex (Figure S2H) and specifically AMPK $\alpha$ 1 bound to recombinant GST-MAGE-A3 and GST-MAGE-A6, but not GST-TRIM28 *in vitro* (Figure 3H). Additionally, overexpressed and endogenous AMPK $\alpha$ 1 co-immunoprecipitated with MAGE-A3 from cells (Figures 3I and S2I). These results suggest that MAGE-A3/6 specifies ubiquitination of AMPK $\alpha$ 1 by the TRIM28 ubiquitin ligase.

### Inhibition of AMPK by MAGE-A3/6-TRIM28 Impacts Cellular Metabolic Processes

Next, we determined whether modulation of AMPK protein levels by MAGE-A3/6-TRIM28 had a functional impact on AMPK activity and the cellular metabolic processes it controls. Knockdown of MAGE-A3/6 or TRIM28 increased both total and active (phosphorylated, pT172) AMPK $\alpha$ 1 (Figure 4A). Furthermore, the knockdown of MAGE-A3/6 or TRIM28 resulted in increased phosphorylation of ACC1 (Figure 4B), a prototypical target of AMPK (Carling et al., 1987). Although AMPK generally antagonizes the Warburg effect (Faubert et al., 2013), acute stimulation of AMPK can promote glycolysis through a variety of activities, including plasma membrane localization of the GLUT1 glucose transporter and increased glucose consumption (Barnes et al., 2002; Hardie et al., 2012b). Knockdown of MAGE-A3/6 or

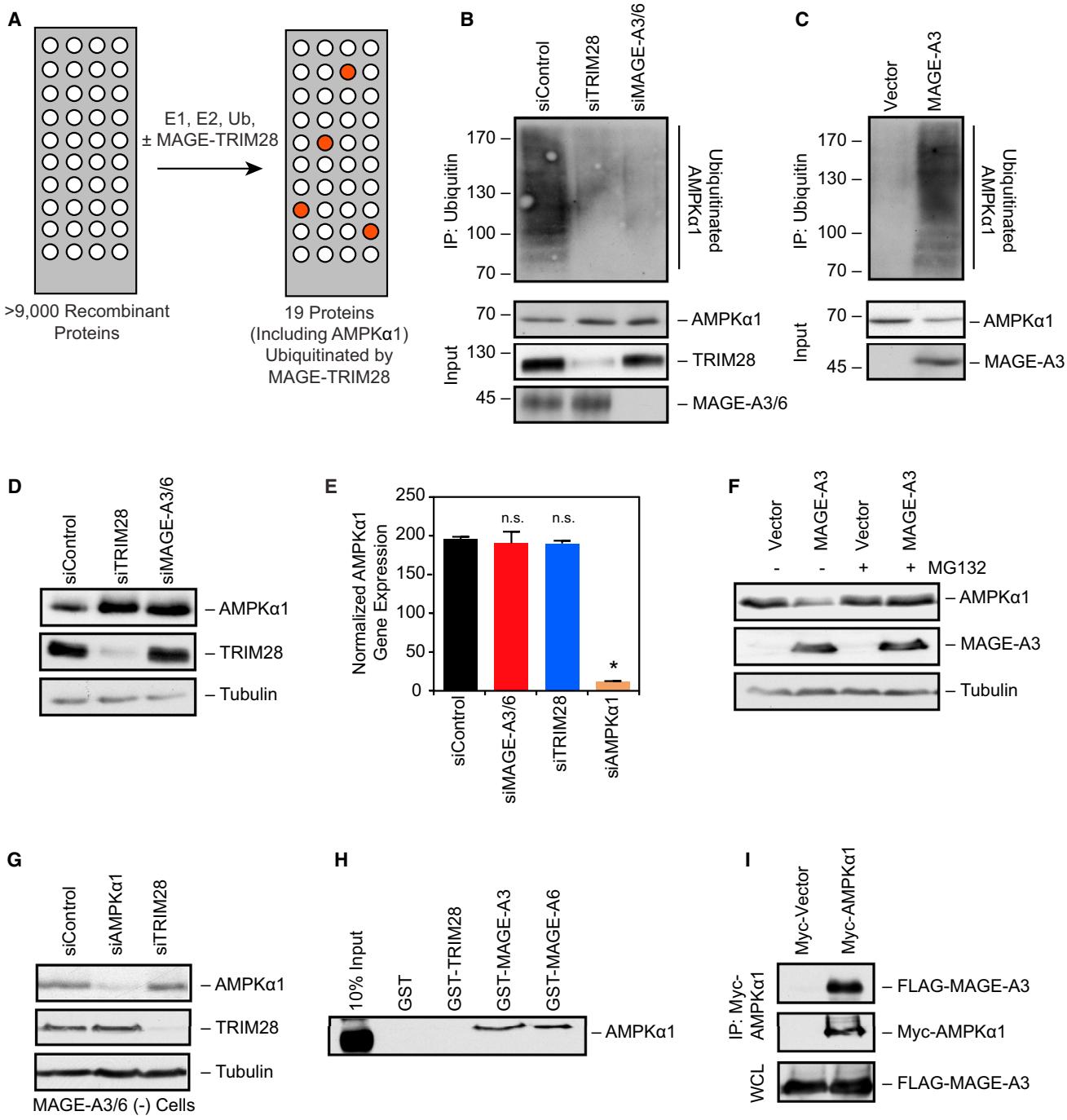
TRIM28 resulted in increased plasma membrane localized GLUT1 (Figure 4C). Furthermore, knockdown of TRIM28 increased glucose consumption (Figure 4D) and correspondingly increased lactate production (Figure 4E) in MAGE-A3/6-positive cells. These results suggest that MAGE-A3/6-TRIM28 has a functional impact on cellular metabolism.

In addition to regulating cellular glucose metabolism, AMPK is well documented to inhibit anabolic processes, such as mTOR-dependent protein synthesis, to conserve energy in the context of reduced ATP levels (Gwinn et al., 2008; Inoki et al., 2003). Therefore, we knocked down MAGE-A3/6 or TRIM28 and followed mTOR activity by examining phosphorylation of p70 ribosomal S6 kinase and ribosomal S6 protein. Upon knockdown of MAGE-A3/6 or TRIM28, mTOR signaling was severely inhibited and phosphorylation of both p70 ribosomal S6 kinase and ribosomal S6 protein were reduced (Figure 4F). Similarly, amino-acid-induced mTOR activity was significantly reduced upon depletion of MAGE-A3/6 (Figure 4G). Importantly, reduction in basal mTOR activity was rescued by treatment with the AMPK inhibitor, compound c (Figure 4H), or co-depletion of AMPK $\alpha$ 1 (Figure 4I). Together, these results suggest that the MAGE-A3/6-TRIM28 ubiquitin ligase is functionally important for maintenance of mTOR activity, likely through inhibition of AMPK.

### MAGE-A3/6-TRIM28 Ubiquitin Ligase Inhibits Autophagy

Because MAGE-A3/6-TRIM28 regulates both AMPK and mTOR activities and both of these signaling pathways converge to opposingly modulate autophagy (Egan et al., 2011; Kim et al., 2011), we examined whether MAGE-A3/6-TRIM28 influences autophagy. One mechanism by which AMPK and mTOR regulate autophagy is through phosphorylation of the proximal ULK1 kinase required for autophagosome formation. AMPK phosphorylation of ULK1 S555 promotes ULK1 activity and autophagy, whereas mTOR phosphorylation of ULK1 S757 inhibits ULK1 activity and autophagy (Egan et al., 2011; Kim et al., 2011). Knockdown of MAGE-A3/6 or TRIM28 upregulated ULK1 S555 phosphorylation (AMPK site) and downregulated ULK1 S757 phosphorylation (mTOR site) (Figure 5A). Changes in ULK1 phosphorylation by MAGE-A3/6 or TRIM28 knockdown were accompanied by the expected increase in GFP-LC3 puncta, a marker of autophagy (Figures 5B and S3A–S3C). The increased GFP-LC3 puncta in MAGE-A3/6- and TRIM28-depleted cells was blocked by co-depletion of ULK1 (Figure 5C). The magnitude of increased GFP-LC3 puncta upon MAGE-A3/6 or TRIM28 knockdown was similar to knockdown of mTOR, an established potent inhibitor of autophagy (Figure 5B). To complement our results using cells stably expressing GFP-LC3, we examined the number of endogenous LC3 puncta upon knockdown of MAGE-A3/6. Similarly to GFP-LC3, siRNAs targeting MAGE-A3/6 induced the accumulation of endogenous LC3 puncta in MAGE-A3/6-positive cells but had no effect in MAGE-A3/6-negative cells (Figures 5D and 5E). Importantly, short-term inhibition of AMPK with compound c attenuated MAGE-A3/6-RNAi-induced GFP-LC3 puncta formation (Figure 5F).

Since an increase in LC3 puncta may represent either a block in autophagosome fusion with lysosomes or an increase in autophagy, we measured the consumption (levels) of GFP-LC3 by



**Figure 3. MAGE-A3/6-TRIM28 E3 Ubiquitin Ligase Ubiquitinates and Degrades AMPK $\alpha$ 1**

(A) Schematic of in vitro screen for MAGE-TRIM28 ubiquitination substrates using protein arrays.

(B) AMPK $\alpha$ 1 ubiquitination requires MAGE-A3/6-TRIM28. HeLa (MAGE-A3/6-positive) were treated with the indicated siRNAs for 24 hr before transfection with Myc-tagged ubiquitin for 48 hr before anti-Myc immunoprecipitation (IP) and immunoblotting was performed ( $n = 3$ ).

(C) Expression of MAGE-A3 promotes AMPK $\alpha$ 1 ubiquitination. MAGE-A3/6-negative HEK293 cells stably expressing FLAG-MAGE-A3 were transfected with Myc-ubiquitin 48 hr before anti-Myc IP and immunoblotting was performed ( $n = 3$ ).

(D) Knockdown of MAGE-A3/6-TRIM28 increases AMPK $\alpha$ 1 protein levels. MAGE-A3/6-positive cells were treated with the indicated siRNAs for 72 hr and then blotted for the indicated proteins ( $n = 3$ ).

(E) Knockdown of MAGE-A3/6-TRIM28 does not affect AMPK $\alpha$ 1 mRNA levels. MAGE-A3/6-positive cells were treated with the indicated siRNAs for 72 hr and then AMPK $\alpha$ 1 mRNA levels were determined by qRT-PCR ( $n = 3$ ). Data are represented as the mean  $\pm$  SD. Asterisks indicate  $p < 0.05$ .

(legend continued on next page)

flow cytometry. We observed a significant decrease in GFP-LC3 fluorescence upon knockdown of MAGE-A3/6 or TRIM28, and this was again similar to the degree of GFP-LC3 consumption upon mTOR depletion (Figures 5G and S3D). These results were further confirmed by western blotting where knockdown of MAGE-A3/6 or TRIM28 promoted a marked decrease in GFP-LC3 protein levels (Figure 5H). These changes in GFP-LC3 were not due to alterations in GFP-LC3 mRNA levels (Figure S3E). To determine if the decrease in LC3 was the consequence of increased autophagic flux, we treated cells depleted of MAGE-A3/6 or TRIM28 with baflomycin A1 to prevent acidification of lysosomes and degradation of proteins by autophagy. Short-term treatment of cells with baflomycin A1 blocked consumption of GFP-LC3 by knockdown of MAGE-A3/6 or TRIM28 (Figure 5I). Finally, we examined the levels of autophagy in MAGE-A3/6- or TRIM28-depleted cells by an independent measure, consumption of the p62/SQSTM1 autophagy adaptor. Similarly to LC3, endogenous p62/SQSTM1 was consumed upon knockdown of MAGE-A3/6 or TRIM28 (Figure 5J), and this could be rescued by baflomycin A1 (Figure 5K). Furthermore, the ability of MAGE-A3/6 to inhibit autophagy was also confirmed by expression of MAGE-A3 in normal, non-transformed cells that typically are negative for MAGE-A3/6. MAGE-A3 expression induced the degradation of AMPK $\alpha$ 1 and the accumulation of p62/SQSTM1 (Figure 5L), consistent with reduced autophagy in these cells. Collectively, these results suggest that MAGE-A3/6-TRIM28 inhibits autophagy and that depletion of MAGE-A3/6 or TRIM28 dramatically increases autophagic flux.

### MAGE-A3/6 Regulation of AMPK $\alpha$ 1 Is Relevant in Human Tumors

Our results suggest that the oncogenic MAGE-A3/6-TRIM28 ubiquitin ligase regulates several cellular metabolic regulatory pathways through ubiquitination and degradation of AMPK $\alpha$ 1. To determine the relevance of these findings to human tumors, we examined whether MAGE-A3/6 expression inversely correlated with AMPK activity and protein levels in patient tumor samples. Indeed, breast invasive carcinoma (Figure S4A), colon adenocarcinoma (Figure 6A), and lung squamous cell carcinoma (Figure 6B) tumors expressing MAGE-A3/6 had significantly reduced total and active (phospho-T172) AMPK $\alpha$  protein levels. This reduction was not a consequence of decreased AMPK $\alpha$ 1 mRNAs in these tumors (Figures 6A, 6B, and S4A). Consistent with these findings, the phosphorylated form of AMPK is downregulated in high proportion of cases of breast cancer (Hadad et al., 2009). In addition, MAGE-A3/6 expression in tumors correlated with reduced downstream AMPK signaling, such as increased markers of mTOR activity (Figure S4B).

Finally, AMPK agonists are of significant interest in treatment and prevention of cancer (Hardie et al., 2012a). Thus, we determined whether AMPK agonists could reverse the phenotypes of MAGE-A3/6 driven anchorage-independent growth and cancer cell viability. The AMPK activating compounds, aminoimidazole carboxamide ribonucleotide (AICAR) and metformin, suppressed the ability of MAGE-A6 to promote anchorage-independent growth of normal HCEC cells and DLD1 colon cancer cells in a dose-dependent manner (Figures S4D–S4F). Importantly, these effects were specific to MAGE-A6-expressing cells as AICAR and metformin minimally affected Apc $^{min}$  or MAGE-B10 driven anchorage-independent growth of HCEC cells (Figures S4D and S4E). Since the cellular effects of both AICAR and metformin extend beyond just activation of AMPK, including affecting mitochondrial respiration (Hardie et al., 2012a), we also examined whether a direct allosteric activator of AMPK, A769662 (Cool et al., 2006; Landgraf et al., 2013), or genetic manipulation of AMPK $\alpha$ 1 could alter phenotypes associated with MAGE-A3/6. MAGE-A6-induced, but not Apc $^{min}$ - or MAGE-B10-induced anchorage-independent growth of HCEC and DLD1 cells was significantly impaired by A769662 in a dose-dependent manner (Figures 6C and S4G). Furthermore, co-depletion of AMPK $\alpha$ 1 rescued MAGE-A3/6-RNAi-induced decrease in cell viability (Figure 6D). Taken together, these results suggest that regulation of AMPK by MAGE-A3/6 is relevant to human tumors and pharmacological agonists of AMPK may have therapeutic potential in MAGE-A3/6-positive tumors.

### DISCUSSION

AMPK senses and responds to the energy status of cells to regulate multiple metabolic processes and limit energy expenditure. Significant effort has been directed toward understanding the role and dysregulation of AMPK in cancer. One known mechanism of reducing AMPK activity in cancer is mutation/deletion of its upstream regulatory kinase Lkb1/Stk11. However, this is a rare event in most tumor types other than lung adenocarcinomas and cervical cancers (Wingo et al., 2009). In this study, we demonstrate that the MAGE-A3/6-TRIM28 E3 ligase complex ubiquitinates and degrades AMPK $\alpha$ 1. Thus, the prominent activation of MAGE-A3/6 expression in many cancer types may represent an alternative mechanism for downregulating the AMPK signaling pathway (Figure S5). Consistent with this, expression of MAGE-A3/6 and mutation of Lkb1/Stk11 are rarely found in the same lung adenocarcinoma tumors (Figure S4C,  $p < 0.01$ ).

MAGE-A3/6 are normally exclusively expressed in the testis but are frequently turned on in many tumor types, including colon, lung, and breast tumors (Figure 1). In combination with

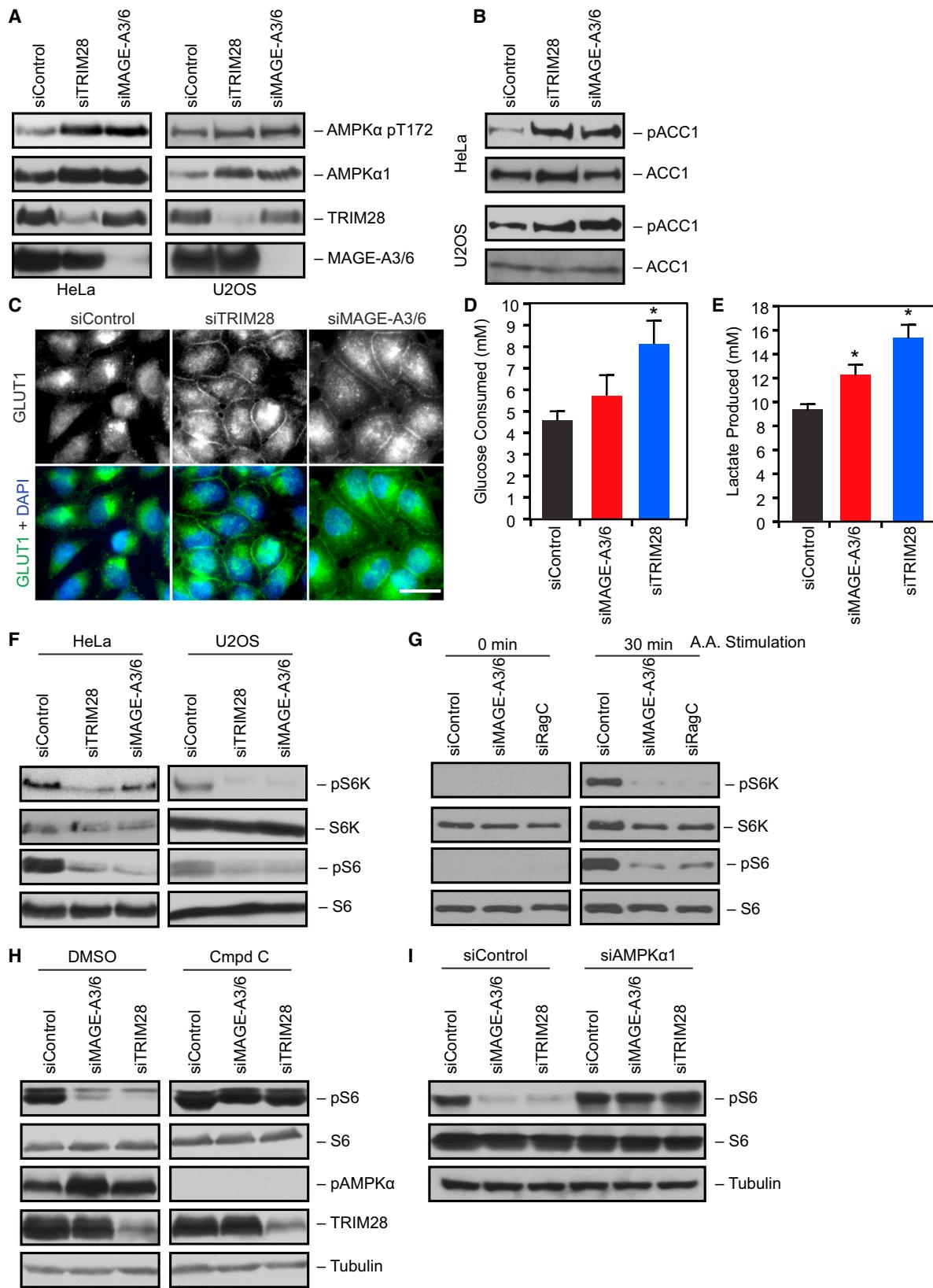
(F) MAGE-A3 promotes proteasome-dependent AMPK $\alpha$ 1 degradation. MAGE-A3/6-negative cells expressing vector of MAGE-A3 were treated with 5  $\mu$ M MG132 for 4 hr before immunoblotting ( $n = 3$ ).

(G) TRIM28-mediated AMPK $\alpha$ 1 degradation requires MAGE-A3/6. MAGE-A3/6-negative HEK293 cells were transfected with the indicated siRNA for 72 hr before cell lysates were immunoblotted ( $n > 3$ ).

(H) GST pull-down assays reveal AMPK $\alpha$ 1 directly binds to MAGE-A3 and MAGE-A6, but not TRIM28 (22-418) or GST ( $n = 3$ ).

(I) HeLa cells expressing FLAG-MAGE-A3 or FLAG-vector along with Myc-AMPK $\alpha$ 1 were subjected to anti-Myc IP and immunoblotting.

WCL, whole-cell lysate. Data are representative of multiple experiments ( $n = 2$ ). See also Table S2.



previous studies, our findings suggest that activation of MAGE-A3/6 in cancer cells is not a by-product, passenger event during cellular transformation and tumorigenesis, but rather MAGE-A3/6 are driver genes that support multiple phenotypes associated with tumors, including metabolic dysregulation. We propose that one critical oncogenic function of MAGE-A3/6 is downregulation of AMPK and alteration of cellular metabolism in cancer cells. Strikingly, this mode of AMPK regulation does not occur in normal somatic cells that do not express MAGE-A3/6 but only occurs upon reactivation of the testicular MAGE-A3/6 program in cancer cells.

Although AMPK coordinates many different actions in the cell, one key process it controls is autophagy. While the role of autophagy in the progression of cancer is multifaceted, loss of autophagy has been implicated in the initiation of tumorigenesis (Choi et al., 2013; Wei et al., 2013; White, 2012). Our results suggest that aberrant activation of MAGE-A3/6 in tumors may provide a unique mechanism for inhibition of tumor-suppressive autophagy during tumor initiation. Interestingly, MAGE-A3/6 expression is undetectable in never-smokers but is aberrantly found in the lungs of smokers before they have any clinical signs of disease (Jang et al., 2001). Thus, MAGE-A3/6 expression may occur early during tumor initiation and could be one mechanism to downregulate autophagy during this stage. Identification of the factors that regulate MAGE-A3/6 expression in adult tissues may provide insights into understanding events leading to tumor initiation. One major regulatory mechanism controlling expression of MAGE cancer-testis antigens is promoter CpG methylation in normal somatic cells (Simpson et al., 2005). However, simple demethylation of MAGEs is not sufficient to drive expression (Weber et al., 1994). The identification of additional transcriptional regulators will be of utmost importance.

Our findings of the association of MAGE-A3/6 expression with AMPK degradation in human tumors has important and potentially immediate implications on the utilization of AMPK activating compounds, such as metformin and A769662, that are vigorously being pursued in the prevention and treatment of cancer (Quinn et al., 2013). While AMPK activating drugs are currently in clinical trials for treatment of a variety of tumor types, the early results thus far have been mixed with no apparent explanation (Quinn et al., 2013). We propose that MAGE-A3/6 expression status may be a useful enrollment biomarker to select patients with the greatest potential response to AMPK agonists. Additionally, since MAGE-A3/6 expression increases signaling through

the mTOR pathway, the use of currently approved mTOR inhibitors may be effective in the future treatment of MAGE-A3/6-driven tumors.

Little is known about the physiological role of MAGE-A3/6 in the testis. Our findings on the molecular and cellular functions of MAGE-A3/6-TRIM28 in cancer provide intriguing insights into their normal physiological function during spermatogenesis. Interestingly, germ cells in the testis switch their carbon energy sources as they differentiate from spermatogonia stem cells to maturing haploid spermatids (Nakamura et al., 1984). We propose that MAGE-A3/6 may function to protect maturing spermatocytes from energy stress by dampening AMPK activation. Also, MAGE-A3/6 might enable developmental stage-dependent activation of anabolic pathways required for normal spermatogenesis, such as lipid and protein synthesis. Furthermore, developing spermatocytes also express an unusual splice variant of LKB1 with a different C-terminal region, which is required for spermiogenesis (Towler et al., 2008). Consistent with these ideas, we have found that mouse MAGE-A genes are highly expressed in pre-pachytene spermatocytes (data not shown) where these regulator events are occurring and previous studies have shown that testis-specific knockout of TRIM28 blocks spermatocyte differentiation, resulting in testis degeneration (Weber et al., 2002).

In summary, our findings illuminate a previously unrecognized, widespread regulation of AMPK during tumorigenesis by a testis-specific ubiquitin ligase, provide an unprecedented molecular mechanism by which MAGE cancer-testis antigens drive tumorigenesis, and have important implications to maximizing the clinical utility of AMPK-directed chemotherapies.

## EXPERIMENTAL PROCEDURES

### Cell Culture and Transfections

Cells were cultured under standard conditions and transfected according to manufacturer's recommendation. Detailed descriptions of cell-culture conditions, transfection procedures, siRNA sequences, and antibodies are described in [Extended Experimental Procedures](#).

### RNA Preparation and qRT-PCR

Preparation of RNA from tissues and cells and qRT-PCR analysis was performed by standard molecular biology techniques and described in [Extended Experimental Procedures](#). All procedures and use of mice were approved by the Institutional Animal Care and Use Committee of UT Southwestern Medical Center.

### Figure 4. MAGE-A3/6-TRIM28 Ubiquitin Ligase Regulates AMPK-Controlled Metabolic Processes

(A and B) MAGE-A3/6-TRIM28 knockdown increases phospho-AMPK (A) and phospho-ACC1 (B) signaling. HeLa or U2OS cells were treated with the indicated siRNAs for 72 hr before cell lysates were immunoblotted (n = 3).

(C) MAGE-A3/6-TRIM28 knockdown increases Glut1 plasma membrane localization. HeLa cells were treated with the indicated siRNAs for 72 hr before immunostaining for Glut1 (n = 3). Scale bar, 20 μm.

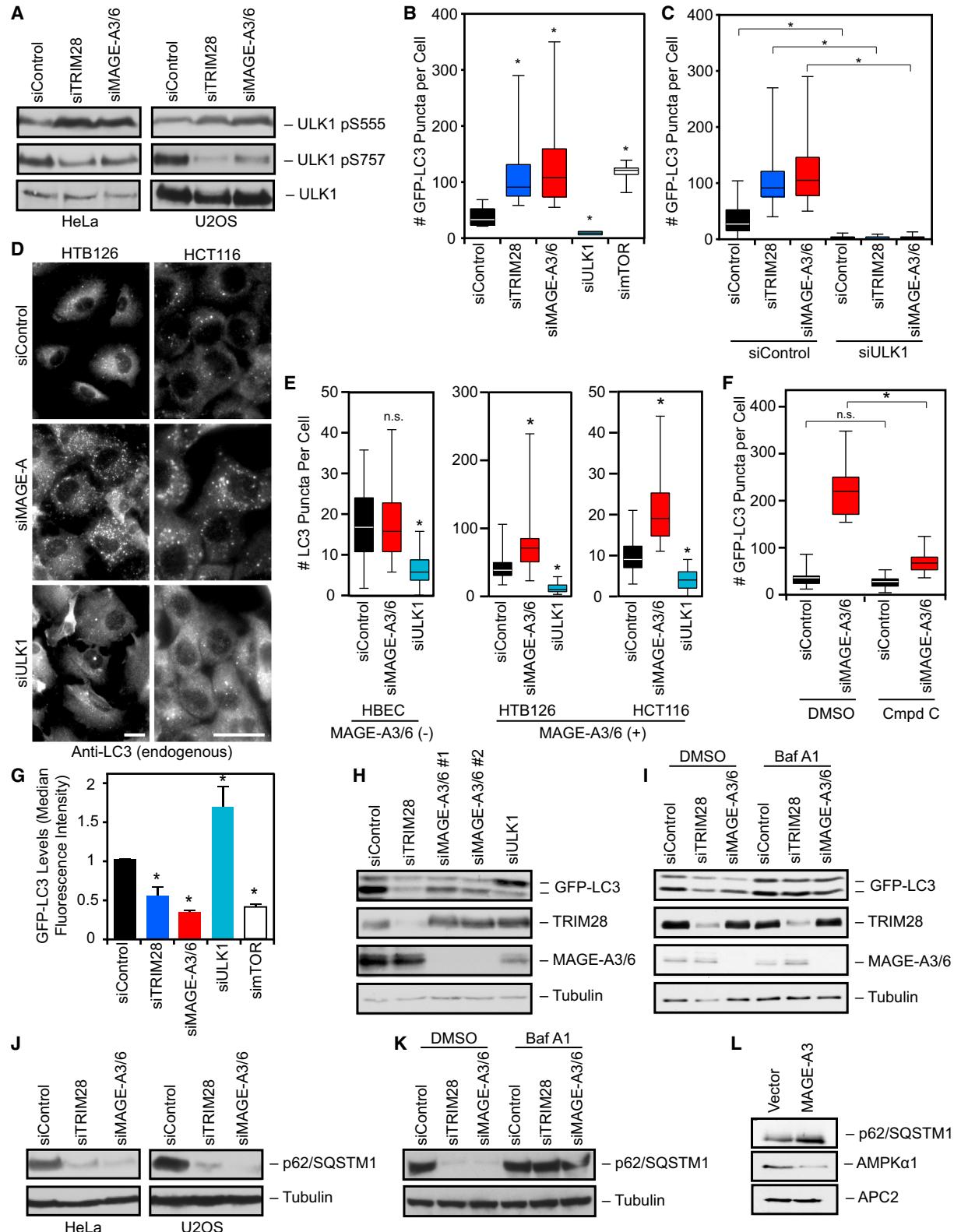
(D and E) MAGE-A3/6-TRIM28 knockdown alters glucose metabolism. HeLa cells were treated with MAGE-A3/6 or TRIM28 siRNA for 72 hr and then fed with fresh media. After 6 hr, media was collected and glucose (D) and lactate (E) levels in media were analyzed via nova analyzer. Data (n = 3) represent mean ± SD.

(F) MAGE-A3/6-TRIM28 is required for mTOR signaling. HeLa or U2OS cells were treated with the indicated siRNAs for 72 hr before immunoblotting (n = 3).

(G) MAGE-A3/6-TRIM28 is required for amino-acid-induced mTOR activity. HeLa cells were treated with siControl, siMAGE-A3/6, or siRagC (positive control) siRNAs for 72 hr before 6 hr starvation in EBSS (0 min) or starvation followed by 30 min amino acid stimulation (n = 3).

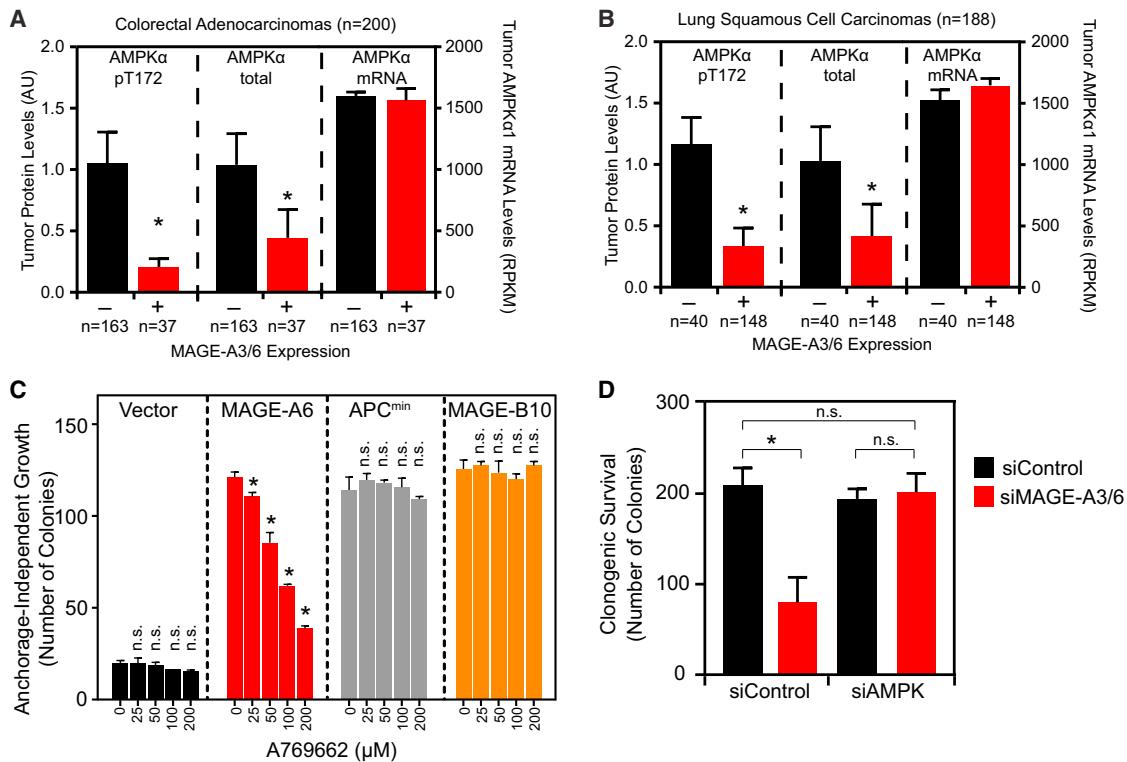
(H and I) Inhibition of AMPK reverses mTOR inhibition by MAGE-A3/6 or TRIM28 knockdown. Cells were transfected with the indicated siRNAs for 72 hr before collection and immunoblotting (I) or treatment for 4 hr with 10 μM compound C or vehicle (DMSO) (H) (n ≥ 2).

Asterisks indicate p < 0.05.

**Figure 5. MAGE-A3/6-TRIM28 Ubiquitin Ligase Inhibits Autophagy**

(A) HeLa or U2OS cells were transfected with siRNA for 72 hr before immunoblotting (ULK1 pS555 AMPK target site; ULK1 pS757 mTOR target site) ( $n > 3$ ).

(legend continued on next page)



**Figure 6. Regulation of AMPK by MAGE-A3/6-TRIM28 Is Relevant to Human Tumors**

(A and B) TCGA data were analyzed for MAGE-A3/6 mRNA levels and total and active (pT172) AMPK $\alpha$  protein levels. Data are mean  $\pm$  SE with number (n) of tumors indicated. Asterisks indicate  $p < 0.01$ .

(C) Anchorage-independent growth assays of the indicated HCEC cells were performed in presence of the indicated concentrations of A769662. Number of colonies  $>100 \mu\text{m}$  were counted (n = 3). Data are mean  $\pm$  SD. Asterisks indicate  $p < 0.05$ .

(D) Colony formation assays were performed in HeLa cells treated with the indicated siRNAs. Data are mean  $\pm$  SD, n = 3. Asterisks indicate  $p < 0.05$ . See also Figures S4 and S5.

#### Colony Formation and Anchorage-Independent Growth Soft Agar Assays

For colony formation assays on plastic, cells were transfected for 72 hr with siRNAs and then re-plated at single-cell density. After 3 weeks, cells were fixed and stained with crystal violet (0.05% [w/v]) and counted. For anchorage-independent growth soft agar growth assays, cells were suspended in 0.375% Noble agar (Difco) supplemented with

regular growth medium and overlaid on 0.5% Noble agar. Cells were allowed to grow for 2–4 weeks before colonies  $\geq 100 \mu\text{m}$  in size were counted.

#### Immunofluorescence and Microscopy

Immunofluorescence was performed essentially as described previously (Hao et al., 2013) and in Extended Experimental Procedures.

- (B) U2OS cells stably expressing GFP-LC3 were treated with the indicated siRNAs for 72 hr before imaging and quantitation of GFP-LC3 puncta (n = 3). Boxplots are mean and quartiles.
- (C) U2OS cells stably expressing GFP-LC3 were treated with the indicated siRNAs for 72 hr before imaging (n = 3). Boxplots represent mean and quartiles of n > 50 cells.
- (D and E) Seventy-two hours after transfection, cells were stained for endogenous LC3 (D). Scale bars, 20  $\mu\text{m}$ . Boxplots of number of LC3 puncta per cell (E). Boxplots are mean and quartiles of n > 50 cells.
- (F) U2OS cells were transfected with the indicated siRNA for 72 hr before treatment for 4 hr with 10  $\mu\text{M}$  compound C or vehicle (DMSO). Boxplots represent mean and quartiles of n > 50 cells.
- (G) U2OS GFP-LC3 cells were transfected with the indicated siRNA for 72 hr. Median GFP fluorescent intensity  $\pm$  SD as determined by flow cytometry is shown (n = 3).
- (H) Cells stably expressing GFP-LC3 were treated with siRNAs for 72 hr before immunoblotting (n > 3).
- (I) Knockdown of MAGE-A3/6 or TRIM28 increases autophagic flux. GFP-LC3 cells were transfected with the indicated siRNA for 72 hr. Cells were treated with baflomycin A for 4 hr before cell lysates were immunoblotted (n = 3).
- (J and K) Knockdown of MAGE-A3/6 or TRIM28 increases p62 consumption. HeLa or U2OS cells were treated with the indicated siRNA for 72 hr. Cell lysates were collected (J) or cells were treated with DMSO or baflomycin A for 4 hr before cell lysates were collected (K), and the indicated proteins were detected by immunoblotting (n  $\geq$  2).
- (L) MAGE-A3/6-negative HBEC cells were stably transfected with Myc-MAGE-A3, and cell lysates were immunoblotted (n = 2).

Asterisks indicate  $p < 0.05$ . See also Figure S3.

**In Vitro Ubiquitination Screen**

ProtoArray containing >9,000 GST-tagged recombinant proteins purified from SF9 insect cells was purchased from Invitrogen. In vitro ubiquitination on the slide was performed according to manufacture instructions with minor modifications described in [Extended Experimental Procedures](#).

**Glucose Consumption and Lactate Measurements**

Twenty-four hours after plating, cells were transfected with siRNA. Seventy-two hours after siRNA transfections, cells were changed into fresh media for 6 hr. Media was collected and analyzed using Nova Analyzer to quantitate amount of glucose and lactate in the media.

**Recombinant Protein Purification and In Vitro Binding Assays**

Recombinant proteins were produced using standard procedures described in [Extended Experimental Procedures](#). In vitro binding assays were performed as described previously (Doyle et al., 2010; Hao et al., 2013) and specified in the [Experimental Procedures](#).

**Assessment of mRNA/Protein Expression Levels in Human Tumors and Statistical Analysis**

mRNA levels, survival data, and mutational status were determined using the cancer genome atlas. Tumor protein expression levels were determined previously by reverse-phase protein arrays performed on tumors with matching RNA sequencing data ([Cancer Genome Atlas Research Network, 2014](#)).

**SUPPLEMENTAL INFORMATION**

Supplemental Information includes Extended Experimental Procedures, five figures, and two tables and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2015.01.034>.

**AUTHOR CONTRIBUTIONS**

C.T.P., S.R., M.B.P., M.A.W., and P.R.P. contributed to experimental design. C.T.P., S.R., K.F.T., J.L.W., M.B.P., Y.O., and P.R.P. performed experiments, data analysis, and figure composition. C.T.P. and P.R.P. wrote the manuscript. S.R., K.F.T., J.L.W., M.B.P., and P.R.P. proofed the manuscript.

**ACKNOWLEDGMENTS**

We thank members of the Potts lab for helpful discussions and critical reading of the manuscript. We also thank Drs. Ralph Deberardinis, Beth Levine, John Minna, Jerry Shay, and Hongtao Yu for guidance and critical reagents. This work was supported by NIH Pharmacological Sciences Training Grant GM007062 (C.T.P.), Michael L. Rosenberg Scholar in Medical Research fund (P.R.P.), CPRIT R1117 (P.R.P.), DOD Discovery Award W81XWH-12-1-0446 (P.R.P.), and WELCH Foundation I-1821 (P.R.P.).

Received: October 8, 2014

Revised: December 3, 2014

Accepted: January 15, 2015

Published: February 12, 2015

**REFERENCES**

Barnes, K., Ingram, J.C., Porras, O.H., Barros, L.F., Hudson, E.R., Fryer, L.G., Foufelle, F., Carling, D., Hardie, D.G., and Baldwin, S.A. (2002). Activation of GLUT1 by metabolic and osmotic stress: potential involvement of AMP-activated protein kinase (AMPK). *J. Cell Sci.* **115**, 2433–2442.

Cancer Genome Atlas Research Network (2014). Comprehensive molecular profiling of lung adenocarcinoma. *Nature* **511**, 543–550, Erratum in *Nature* 2014;514:262. Rogers, K [corrected to Rodgers, K].

Cantó, C., Gerhart-Hines, Z., Feige, J.N., Lagouge, M., Noriega, L., Milne, J.C., Elliott, P.J., Puigserver, P., and Auwerx, J. (2009). AMPK regulates energy expenditure by modulating NAD<sup>+</sup> metabolism and SIRT1 activity. *Nature* **458**, 1056–1060.

Carling, D., Zammit, V.A., and Hardie, D.G. (1987). A common bicyclic protein kinase cascade inactivates the regulatory enzymes of fatty acid and cholesterol biosynthesis. *FEBS Lett.* **223**, 217–222.

Choi, A.M., Ryter, S.W., and Levine, B. (2013). Autophagy in human health and disease. *N. Engl. J. Med.* **368**, 651–662.

Chomez, P., De Backer, O., Bertrand, M., De Plaein, E., Boon, T., and Lucas, S. (2001). An overview of the MAGE gene family with the identification of all human members of the family. *Cancer Res.* **61**, 5544–5551.

Chou, C.C., Lee, K.H., Lai, I.L., Wang, D., Mo, X., Kulp, S.K., Shapiro, C.L., and Chen, C.S. (2014). AMPK reverses the mesenchymal phenotype of cancer cells by targeting the Akt-MDM2-Foxo3a signaling axis. *Cancer Res.* **74**, 4783–4795.

Cool, B., Zinker, B., Chiou, W., Kifle, L., Cao, N., Perham, M., Dickinson, R., Adler, A., Gagne, G., Iyengar, R., et al. (2006). Identification and characterization of a small molecule AMPK activator that treats key components of type 2 diabetes and the metabolic syndrome. *Cell Metab.* **3**, 403–416.

De Plaein, E., Arden, K., Traversari, C., Gaforio, J.J., Szikora, J.P., De Smet, C., Brasseur, F., van der Bruggen, P., Lethé, B., Lurquin, C., et al. (1994). Structure, chromosomal localization, and expression of 12 genes of the MAGE family. *Immunogenetics* **40**, 360–369.

Doyle, J.M., Gao, J., Wang, J., Yang, M., and Potts, P.R. (2010). MAGE-RING protein complexes comprise a family of E3 ubiquitin ligases. *Mol. Cell* **39**, 963–974.

Egan, D.F., Shackelford, D.B., Mihaylova, M.M., Gelino, S., Kohnz, R.A., Mair, W., Vasquez, D.S., Joshi, A., Gwinn, D.M., Taylor, R., et al. (2011). Phosphorylation of ULK1 (hATG1) by AMP-activated protein kinase connects energy sensing to mitophagy. *Science* **331**, 456–461.

Esteve-Puig, R., Canals, F., Colomé, N., Merlino, G., and Recio, J.A. (2009). Uncoupling of the LKB1-AMPKalpha energy sensor pathway by growth factors and oncogenic BRAF. *PLoS ONE* **4**, e4771.

Faubert, B., Boily, G., Izreig, S., Griss, T., Samborska, B., Dong, Z., Dupuy, F., Chambers, C., Fuerth, B.J., Viollet, B., et al. (2013). AMPK is a negative regulator of the Warburg effect and suppresses tumor growth in vivo. *Cell Metab.* **17**, 113–124.

Feng, Y., Gao, J., and Yang, M. (2011). When MAGE meets RING: insights into biological functions of MAGE proteins. *Protein Cell* **2**, 7–12.

Gwinn, D.M., Shackelford, D.B., Egan, D.F., Mihaylova, M.M., Mery, A., Vasquez, D.S., Turk, B.E., and Shaw, R.J. (2008). AMPK phosphorylation of raptor mediates a metabolic checkpoint. *Mol. Cell* **30**, 214–226.

Hadad, S.M., Baker, L., Quinlan, P.R., Robertson, K.E., Bray, S.E., Thomson, G., Kellock, D., Jordan, L.B., Purdie, C.A., Hardie, D.G., et al. (2009). Histological evaluation of AMPK signalling in primary breast cancer. *BMC Cancer* **9**, 307.

Hadad, S., Iwamoto, T., Jordan, L., Purdie, C., Bray, S., Baker, L., Jellema, G., Deharo, S., Hardie, D.G., Pusztai, L., et al. (2011). Evidence for biological effects of metformin in operable breast cancer: a pre-operative, window-of-opportunity, randomized trial. *Breast Cancer Res. Treat.* **128**, 783–794.

Hao, Y.H., Doyle, J.M., Ramanathan, S., Gomez, T.S., Jia, D., Xu, M., Chen, Z.J., Billadeau, D.D., Rosen, M.K., and Potts, P.R. (2013). Regulation of WASH-dependent actin polymerization and protein trafficking by ubiquitination. *Cell* **152**, 1051–1064.

Hardie, D.G., and Alessi, D.R. (2013). LKB1 and AMPK and the cancer-metabolism link – ten years after. *BMC Biol.* **11**, 36.

Hardie, D.G., Ross, F.A., and Hawley, S.A. (2012a). AMP-activated protein kinase: a target for drugs both ancient and modern. *Chem. Biol.* **19**, 1222–1236.

Hardie, D.G., Ross, F.A., and Hawley, S.A. (2012b). AMPK: a nutrient and energy sensor that maintains energy homeostasis. *Nat. Rev. Mol. Cell Biol.* **13**, 251–262.

Hawley, S.A., Boudeau, J., Reid, J.L., Mustard, K.J., Udd, L., Mäkelä, T.P., Alessi, D.R., and Hardie, D.G. (2003). Complexes between the LKB1 tumor suppressor, STRAD alpha/beta and MO25 alpha/beta are upstream kinases in the AMP-activated protein kinase cascade. *J. Biol.* **2**, 28.

Hawley, S.A., Pan, D.A., Mustard, K.J., Ross, L., Bain, J., Edelman, A.M., Frenquelli, B.G., and Hardie, D.G. (2005). Calmodulin-dependent protein kinase kinase-beta is an alternative upstream kinase for AMP-activated protein kinase. *Cell Metab.* 2, 9–19.

Imamura, K., Ogura, T., Kishimoto, A., Kaminishi, M., and Esumi, H. (2001). Cell cycle regulation via p53 phosphorylation by a 5'-AMP activated protein kinase activator, 5-aminoimidazole-4-carboxamide-1-beta-D-ribofuranoside, in a human hepatocellular carcinoma cell line. *Biochem. Biophys. Res. Commun.* 287, 562–567.

Inoki, K., Zhu, T., and Guan, K.L. (2003). TSC2 mediates cellular energy response to control cell growth and survival. *Cell* 115, 577–590.

Jang, S.J., Soria, J.C., Wang, L., Hassan, K.A., Morice, R.C., Walsh, G.L., Hong, W.K., and Mao, L. (2001). Activation of melanoma antigen tumor antigens occurs early in lung carcinogenesis. *Cancer Res.* 61, 7959–7963.

Jones, R.G., Plas, D.R., Kubek, S., Buzzai, M., Mu, J., Xu, Y., Birnbaum, M.J., and Thompson, C.B. (2005). AMP-activated protein kinase induces a p53-dependent metabolic checkpoint. *Mol. Cell* 18, 283–293.

Kim, J., Kundu, M., Viollet, B., and Guan, K.L. (2011). AMPK and mTOR regulate autophagy through direct phosphorylation of Ulk1. *Nat. Cell Biol.* 13, 132–141.

Kim, J., Kim, Y.C., Fang, C., Russell, R.C., Kim, J.H., Fan, W., Liu, R., Zhong, Q., and Guan, K.L. (2013). Differential regulation of distinct Vps34 complexes by AMPK in nutrient stress and autophagy. *Cell* 152, 290–303.

Laderoute, K.R., Calaoagan, J.M., Chao, W.R., Dinh, D., Denko, N., Duellman, S., Kalra, J., Liu, X., Papandreou, I., Sambucetti, L., and Boros, L.G. (2014). 5'-AMP-activated protein kinase (AMPK) supports the growth of aggressive experimental human breast cancer tumors. *J. Biol. Chem.* 289, 22850–22864.

Landgraf, R.R., Goswami, D., Rajamohan, F., Harris, M.S., Calabrese, M.F., Hoth, L.R., Magyar, R., Pascal, B.D., Chalmers, M.J., Busby, S.A., et al. (2013). Activation of AMP-activated protein kinase revealed by hydrogen/deuterium exchange mass spectrometry. *Structure* 21, 1942–1953.

Lee, C.W., Wong, L.L., Tse, E.Y., Liu, H.F., Leong, V.Y., Lee, J.M., Hardie, D.G., Ng, I.O., and Ching, Y.P. (2012). AMPK promotes p53 acetylation via phosphorylation and inactivation of SIRT1 in liver cancer cells. *Cancer Res.* 72, 4394–4404.

Liang, J., Shao, S.H., Xu, Z.X., Hennessy, B., Ding, Z., Larrea, M., Kondo, S., Dumont, D.J., Guterman, J.U., Walker, C.L., et al. (2007). The energy sensing LKB1-AMPK pathway regulates p27(kip1) phosphorylation mediating the decision to enter autophagy or apoptosis. *Nat. Cell Biol.* 9, 218–224.

Liu, W., Cheng, S., Asa, S.L., and Ezzat, S. (2008). The melanoma-associated antigen A3 mediates fibronectin-controlled cancer progression and metastasis. *Cancer Res.* 68, 8104–8112.

Marcar, L., MacLaine, N.J., Hupp, T.R., and Meek, D.W. (2010). Mage-A cancer/testis antigens inhibit p53 function by blocking its interaction with chromatin. *Cancer Res.* 70, 10362–10370.

Matsumoto, S., Iwakawa, R., Takahashi, K., Kohno, T., Nakanishi, Y., Matsuno, Y., Suzuki, K., Nakamoto, M., Shimizu, E., Minna, J.D., and Yokota, J. (2007). Prevalence and specificity of LKB1 genetic alterations in lung cancers. *Oncogene* 26, 5911–5918.

Monte, M., Simonatto, M., Peche, L.Y., Bublik, D.R., Gobessi, S., Pierotti, M.A., Rodolfo, M., and Schneider, C. (2006). MAGE-A tumor antigens target p53 transactivation function through histone deacetylase recruitment and confer resistance to chemotherapeutic agents. *Proc. Natl. Acad. Sci. USA* 103, 11160–11165.

Nakamura, M., Okinaga, S., and Arai, K. (1984). Metabolism of round spermatids: evidence that lactate is preferred substrate. *Am. J. Physiol.* 247, E234–E242.

Niraula, S., Dowling, R.J., Ennis, M., Chang, M.C., Done, S.J., Hood, N., Escalante, J., Leong, W.L., McCready, D.R., Reedijk, M., et al. (2012). Metformin in early breast cancer: a prospective window of opportunity neoadjuvant study. *Breast Cancer Res. Treat.* 135, 821–830.

Pernicova, I., and Korbonits, M. (2014). Metformin—mode of action and clinical implications for diabetes and cancer. *Nat. Rev. Endocrinol.* 10, 143–156.

Quinn, B.J., Kitagawa, H., Memmott, R.M., Gills, J.J., and Dennis, P.A. (2013). Repositioning metformin for cancer prevention and treatment. *Trends Endocrinol. Metab.* 24, 469–480.

Roig, A.I., Eskiocak, U., Hight, S.K., Kim, S.B., Delgado, O., Souza, R.F., Spechler, S.J., Wright, W.E., and Shay, J.W. (2010). Immortalized epithelial cells derived from human colon biopsies express stem cell markers and differentiate in vitro. *Gastroenterology* 138, 1012–1021, e1011–1015.

Sanchez-Cespedes, M., Parrella, P., Esteller, M., Nomoto, S., Trink, B., Engles, J.M., Westra, W.H., Herman, J.G., and Sidransky, D. (2002). Inactivation of LKB1/STK11 is a common event in adenocarcinomas of the lung. *Cancer Res.* 62, 3659–3662.

Shackelford, D.B., and Shaw, R.J. (2009). The LKB1-AMPK pathway: metabolism and growth control in tumour suppression. *Nat. Rev. Cancer* 9, 563–575.

Shackelford, D.B., Vasquez, D.S., Corbeil, J., Wu, S., Leblanc, M., Wu, C.L., Vera, D.R., and Shaw, R.J. (2009). mTOR and HIF-1alpha-mediated tumor metabolism in an LKB1 mouse model of Peutz-Jeghers syndrome. *Proc. Natl. Acad. Sci. USA* 106, 11137–11142.

Shantha Kumara, H.M., Grieco, M.J., Caballero, O.L., Su, T., Ahmed, A., Ritter, E., Gnjatic, S., Cekic, V., Old, L.J., Simpson, A.J., et al. (2012). MAGE-A3 is highly expressed in a subset of colorectal cancer patients. *Cancer Immun.* 12, 16.

Shaw, R.J., Kosmatka, M., Bardeesy, N., Hurley, R.L., Witters, L.A., DePinho, R.A., and Cantley, L.C. (2004). The tumor suppressor LKB1 kinase directly activates AMP-activated kinase and regulates apoptosis in response to energy stress. *Proc. Natl. Acad. Sci. USA* 101, 3329–3335.

Simpson, A.J., Caballero, O.L., Jungbluth, A., Chen, Y.T., and Old, L.J. (2005). Cancer/testis antigens, gametogenesis and cancer. *Nat. Rev. Cancer* 5, 615–625.

Suter, M., Riek, U., Tuerk, R., Schlattner, U., Wallimann, T., and Neumann, D. (2006). Dissecting the role of 5'-AMP for allosteric stimulation, activation, and deactivation of AMP-activated protein kinase. *J. Biol. Chem.* 281, 32207–32216.

Towler, M.C., Fogarty, S., Hawley, S.A., Pan, D.A., Martin, D.M., Morrice, N.A., McCarthy, A., Galardo, M.N., Meroni, S.B., Cigorraga, S.B., et al. (2008). A novel short splice variant of the tumour suppressor LKB1 is required for spermiogenesis. *Biochem. J.* 416, 1–14.

Wang, C., Ivanov, A., Chen, L., Fredericks, W.J., Seto, E., Rauscher, F.J., 3rd, and Chen, J. (2005). MDM2 interaction with nuclear corepressor KAP1 contributes to p53 inactivation. *EMBO J.* 24, 3279–3290.

Weber, J., Salgaller, M., Samid, D., Johnson, B., Herlyn, M., Lassam, N., Treisman, J., and Rosenberg, S.A. (1994). Expression of the MAGE-1 tumor antigen is up-regulated by the demethylating agent 5-aza-2'-deoxycytidine. *Cancer Res.* 54, 1766–1771.

Weber, P., Cammas, F., Gerard, C., Metzger, D., Champon, P., Loisson, R., and Mark, M. (2002). Germ cell expression of the transcriptional co-repressor TIF1-beta is required for the maintenance of spermatogenesis in the mouse. *Development* 129, 2329–2337.

Wei, Y., Zou, Z., Becker, N., Anderson, M., Sumpter, R., Xiao, G., Kinch, L., Kordur, P., Christudass, C.S., Veltri, R.W., et al. (2013). EGFR-mediated Beclin 1 phosphorylation in autophagy suppression, tumor progression, and tumor chemoresistance. *Cell* 154, 1269–1284.

Weinstein, I.B. (2002). Cancer. Addiction to oncogenes—the Achilles heel of cancer. *Science* 297, 63–64.

White, E. (2012). Deconvoluting the context-dependent role for autophagy in cancer. *Nat. Rev. Cancer* 12, 401–410.

Winder, W.W., Holmes, B.F., Rubink, D.S., Jensen, E.B., Chen, M., and Hollloszy, J.O. (2000). Activation of AMP-activated protein kinase increases mitochondrial enzymes in skeletal muscle. *J. Appl. Physiol.* 88, 2219–2226.

Wingo, S.N., Gallardo, T.D., Akbay, E.A., Liang, M.C., Contreras, C.M., Boren, T., Shimamura, T., Miller, D.S., Sharpless, N.E., Bardeesy, N., et al. (2009). Somatic LKB1 mutations promote cervical cancer progression. *PLoS ONE* 4, e5137.

Woods, A., Dickerson, K., Heath, R., Hong, S.P., Momcilovic, M., Johnstone, S.R., Carlson, M., and Carling, D. (2005). Ca<sup>2+</sup>/calmodulin-dependent protein kinase kinase-beta acts upstream of AMP-activated protein kinase in mammalian cells. *Cell Metab.* 2, 21–33.

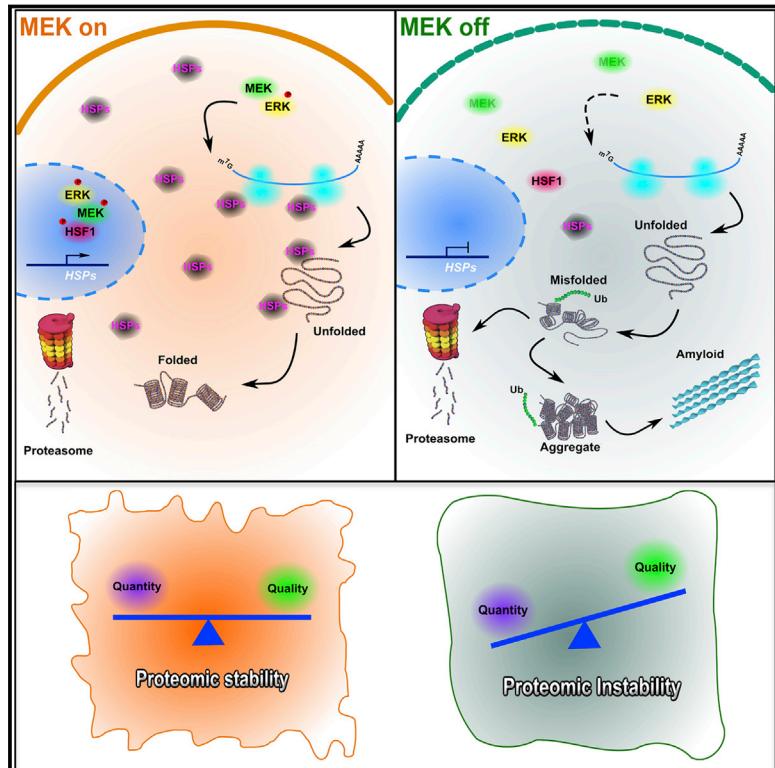
Yang, B., O'Herrin, S.M., Wu, J., Reagan-Shaw, S., Ma, Y., Bhat, K.M., Gravkamp, C., Setaluri, V., Peters, N., Hoffmann, F.M., et al. (2007). MAGE-A, mMage-b, and MAGE-C proteins form complexes with KAP1 and suppress p53-dependent apoptosis in MAGE-positive cell lines. *Cancer Res.* 67, 9954–9962.

Zheng, B., Jeong, J.H., Asara, J.M., Yuan, Y.Y., Granter, S.R., Chin, L., and Cantley, L.C. (2009). Oncogenic B-RAF negatively regulates the tumor suppressor LKB1 to promote melanoma cell proliferation. *Mol. Cell* 33, 237–247.

Zheng, L., Yang, W., Wu, F., Wang, C., Yu, L., Tang, L., Qiu, B., Li, Y., Guo, L., Wu, M., et al. (2013). Prognostic significance of AMPK activation and therapeutic effects of metformin in hepatocellular carcinoma. *Clin. Cancer Res.* 19, 5372–5380.

# MEK Guards Proteome Stability and Inhibits Tumor-Suppressive Amyloidogenesis via HSF1

## Graphical Abstract



## Authors

Zijian Tang, Siyuan Dai, ...,  
Stephen Byers Sampson, Chengkai Dai

## Correspondence

[chengkai.dai@jax.org](mailto:chengkai.dai@jax.org)

## In Brief

The identification of HSF1 as a new substrate for MEK reveals that the RAS/MAP kinase pathway regulates proteostasis in normal cells and that it can be targeted to promote proteomic instability and amyloidogenesis in cancer cells.

## Highlights

- MEK physically interacts with and phosphorylates HSF1 at Ser326
- MEK mobilizes and ERK suppresses the HSF1-mediated proteotoxic stress response
- MEK inhibition disrupts proteostasis and provokes tumor-suppressive amyloidogenesis
- Malignancy is susceptible to proteomic perturbation and subsequent amyloidogenesis

# MEK Guards Proteome Stability and Inhibits Tumor-Suppressive Amyloidogenesis via HSF1

Zijian Tang,<sup>1,2</sup> Siyuan Dai,<sup>1</sup> Yishu He,<sup>1</sup> Rosalinda A. Doty,<sup>1</sup> Leonard D. Shultz,<sup>1</sup> Stephen Byers Sampson,<sup>1</sup> and Chengkai Dai<sup>1,\*</sup>

<sup>1</sup>The Jackson Laboratory, 600 Main Street, Bar Harbor, ME 04609, USA

<sup>2</sup>Graduate Programs, Department of Molecular and Biomedical Sciences, The University of Maine, 5735 Hitchner Hall, Orono, ME 04469, USA

\*Correspondence: chengkai.dai@jax.org

<http://dx.doi.org/10.1016/j.cell.2015.01.028>

## SUMMARY

Signaling through RAS/MAP kinase pathway is central to biology. ERK has long been perceived as the only substrate for MEK. Here, we report that HSF1, the master regulator of the proteotoxic stress response, is a new MEK substrate. Beyond mediating cell-environment interactions, the MEK-HSF1 regulation impacts malignancy. In tumor cells, MEK blockade inactivates HSF1 and thereby provokes proteomic chaos, presented as protein destabilization, aggregation, and, strikingly, amyloidogenesis. Unlike their non-transformed counterparts, tumor cells are particularly susceptible to proteomic perturbation and amyloid induction. Amyloidogenesis is tumor suppressive, reducing *in vivo* melanoma growth and contributing to the potent anti-neoplastic effects of proteotoxic stressors. Our findings unveil a key biological function of the oncogenic RAS-MEK signaling in guarding proteostasis and suppressing amyloidogenesis. Thus, proteomic instability is an intrinsic feature of malignant state, and disrupting the fragile tumor proteostasis to promote amyloidogenesis may be a feasible therapeutic strategy.

## INTRODUCTION

Following environmental challenges, cells stimulate production of heat-shock proteins (HSPs). This HSP induction is the hallmark of the heat-shock, or proteotoxic stress, response (PSR) (Lindquist, 1986). As molecular chaperones, HSPs facilitate folding, transportation, and degradation of other proteins (Morimoto, 2008). In guarding the proteome against misfolding and aggregation, the PSR preserves proteostasis (Balch et al., 2008).

In vertebrates, heat shock transcription factors (HSFs) govern the PSR. Among them is HSF1, the master regulator of this response (Morimoto, 2008; Xiao et al., 1999). As a multi-step process, HSF1 activation entails trimerization, nuclear translocation, posttranslational modifications, and DNA binding (Morimoto, 2008). Yet, our understanding of this process remains incomplete.

The HSF1-mediated PSR antagonizes many pathological conditions, including hyperthermia, heavy-metal toxification,

ischemia, and reperfusion, and oxidative damage, and impacts aging and neurodegeneration (Dai et al., 2012a). HSF1, not surprisingly, acts as a longevity factor (Hsu et al., 2003). In contrast, our and others' work has revealed a pro-oncogenic role of HSF1 (Dai et al., 2007, 2012b; Jin et al., 2011; Meng et al., 2010; Min et al., 2007). Despite its dispensability under non-stress conditions, HSF1 is crucial for tumor cells' growth and survival (Dai et al., 2007). Nonetheless, the mechanisms underlying its activation in malignancy remain unclear.

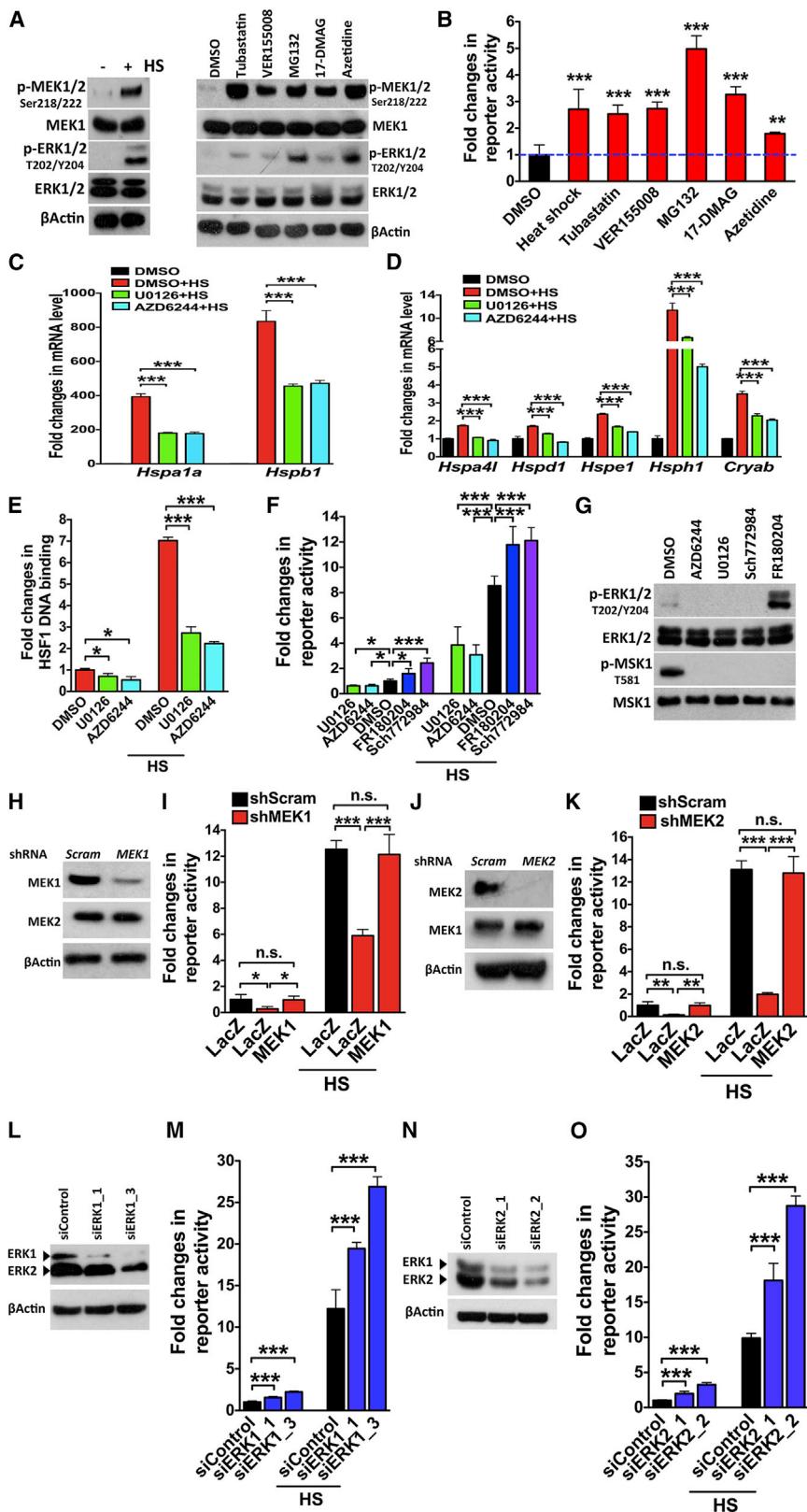
Here, we report that RAS-MEK-ERK signaling critically regulates the PSR. It is MEK that phosphorylates and activates HSF1. MEK inhibition destabilizes the proteome, provoking protein aggregation and amyloidogenesis. Combinatorial proteasome blockade potently augments this tumor-suppressive amyloidogenic effect. Hence, our findings not only suggest HSF1 as a new MEK substrate but also uncover a biological function of RAS-MEK-ERK signaling in governing proteostasis. Beyond shifting the canonical view of RAS-MEK-ERK signaling, our proof-of-concept experiments suggest that intrinsic proteomic instability associated with malignant state may be exploited to combat cancer.

## RESULTS

### MEK and ERK Inversely Regulate the PSR

Phosphorylation notably impacts HSF1 activation (Guettouche et al., 2005), suggesting a key role of signaling pathways. To illuminate how such pathways regulate the PSR, we first examined their responses to stress, focusing on RAS-MEK-ERK signaling. To inflict proteotoxic stress, we applied stressors with diverse mechanisms of action, including heat shock (HS), proteasome inhibitor MG132, histone deacetylase 6 inhibitor tubastatin, amino-acid analog azetidine, and HSP inhibitors (17-DMAG for HSP90 and VER155008 for HSP70) (Kawaguchi et al., 2003; Massey et al., 2010; Morimoto, 2008; Neckers and Workman, 2012). Transient exposure to stressors did not impair cell viability (Figure S1A), but elevated phosphorylation of MEK and ERK (Figure 1A), two key components of this pathway. MEK Ser218/222 and ERK Thr202/Tyr204 phosphorylation signify their active state (Dhillon et al., 2007; Roux and Blenis, 2004). Congruently, all stressors activated ELK1 (Figure 1B), a transcription factor downstream of ERK (Roux and Blenis, 2004).

To determine whether MEK-ERK signaling regulates the PSR, we employed U0126 and AZD6244, two specific MEK1/2



**Figure 1. MEK and ERK Oppositely Regulate the PSR**

(A) NIH 3T3 cells were treated with HS at 43°C for 30 min, 10 μM tubastatin A for 5 hr, 40 μM VER155008 for 1 hr, 500 nM MG132 for 1 hr, 200 nM 17-DMAG for 1 hr, and 2.5 mM azetidine for 15 min.

(B) The dual ELK1 reporter system, comprising a serum response element (SRE)-driven secreted embryonic alkaline phosphatase (SEAP) plasmid and a CMV-driven Gaussia luciferase (GLuc) plasmid, was transfected into HEK293T cells. After 24 hr, cells were treated as in (A) and recovered overnight before measuring reporter activities (mean ± SD, n = 6, ANOVA).

(C and D) NIH 3T3 cells were treated with 20 μM U0126 or 20 nM AZD6244 for 3 hr followed by HS and 4 hr recovery. mRNA levels were quantitated by qRT-PCR (mean ± SD, n = 3, Student's t test).

(E) Immediately after HS, nuclear proteins of NIH 3T3 cells treated as in (C) were extracted to measure HSF1-DNA binding by an ELISA-based assay (mean ± SD, n = 3, ANOVA).

(F) HEK293T cells were transfected with dual HSF1 reporter plasmids, a heat-shock element (HSE)-driven SEAP plasmid and a CMV-GLuc plasmid. After 24 hr, cells were treated with 20 μM U0126, 20 nM AZD6244, 1 μM FR180204, or 100 nM Sch772984 for 3 hr followed by HS for 30 min and overnight recovery (mean ± SD, n = 6, ANOVA).

(G) HEK293T cells were treated with different inhibitors overnight.

(H–K) A LacZ or MEK isoform plasmid was cotransfected with dual HSF1 reporter plasmids into HEK293T cells transduced with lentiviral small hairpin RNAs (shRNAs). After 24 hr, cells were heat shocked at 43°C for 30 min followed by overnight recovery (mean ± SD, n = 3, ANOVA).

(L–O) HEK293T cells were transfected with small interfering RNAs (siRNAs) for 48 hr followed by transfection with dual HSF1 reporter plasmids for 24 hr before HS (mean ± SD, n = 6, ANOVA).

See also Figure S1.

inhibitors (Favata et al., 1998; Yeh et al., 2007). Both inhibitors impeded the HS-induced transcription of *Hsp* genes and impaired the DNA-binding capacity and transcriptional activation of HSF1 (Figures 1C–1F and S1B), suggesting that MEK activates the HSF1-mediated PSR. ERK, phosphorylated by MEK (Ahn et al., 1991), is widely recognized as the master effector of this pathway (Dhillon et al., 2007; Roux and Blenis, 2004). Surprisingly, ERK inhibitors, FR180204 and Sch772984 (Ohori et al., 2005; Morris et al., 2013), activated HSF1 (Figures 1F and S1C). Both MEK and ERK inhibitors impaired two ERK-mediated events—MSK1 phosphorylation and ELK1 activation (Figures 1G and S1D) (Roux and Blenis, 2004). While MEK inhibitors reduced ERK phosphorylation, two ERK inhibitors showed distinct effects (Figure 1G). Sch772984 suppressed ERK phosphorylation, likely due to ERK conformational changes that block MEK-mediated phosphorylation (Morris et al., 2013); conversely, FR180204 promoted ERK phosphorylation (Figure 1G), suggesting feedback MEK activation.

The impacts of MEK and ERK inhibitors on HSF1 were validated via genetic depletions of *MEK* and *ERK* (Figures 1I–1O). While depletion of one *ERK* isoform diminished the other isoform at the protein level (Figures 1L and 1N), mRNA levels of the isoform not targeted were elevated (Figures S1E–S1H), suggesting posttranscriptional mechanisms underlying reduced proteins. These results not only pinpoint RAS-MEK-ERK signaling as a key regulator of the PSR, but also reveal divergent impacts of MEK and ERK on HSF1.

### MEK Physically Interacts with HSF1

To determine whether MEK directly activates HSF1, we examined endogenous MEK-HSF1 interactions by co-immunoprecipitation (coIP). While no evident MEK1/2 proteins were precipitated with HSF1 without HS, HS caused a marked coIP (Figures 2A and 2B), showing a stress-inducible MEK-HSF1 interaction. The mobility shift of HSF1 marks HS-induced phosphorylation (Figure 2A). These MEK-HSF1 interactions were verified via expression of recombinant proteins (Figures S2A and S2B). To determine whether MEK and HSF1 are in direct contact, we employed the Proximity Ligation Assay (PLA) technique (Clausson et al., 2011). Antibody specificities were validated by immunostaining (Figures S2C and S2D). In *MEK*-proficient cells, PLA signals were marginally visible without HS and HS intensified these signals (Figure 2C). In *MEK*-deficient cells, only faint signals were detected even after HS (Figure 2C), confirming the specificity of PLA. Of note, PLA signals were more manifest in the nucleus than in the cytoplasm (Figure 2C), revealing a predominantly nuclear localization of interactions. These results strongly suggest a direct MEK-HSF1 association.

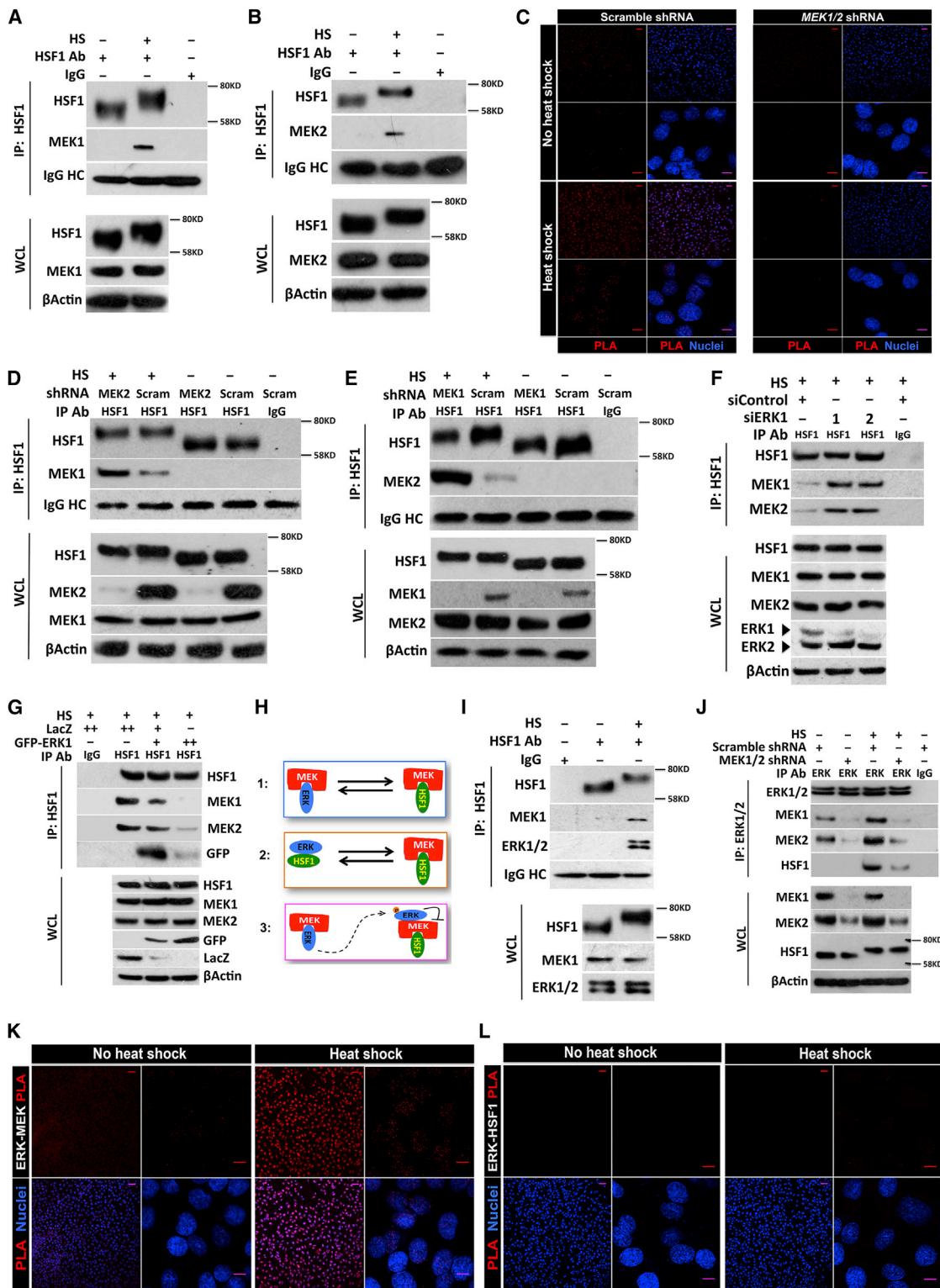
MEK1 and MEK2 form either homo- or heterodimers *in vivo* (Catalanotti et al., 2009). To address which type of dimer binds HSF1, we examined MEK1-HSF1 interactions in the deficiency of MEK2. Under HS, more MEK1 proteins were precipitated with HSF1 in *MEK2*-deficient cells (Figure 2D). Similarly, *MEK1* deficiency heightened MEK2-HSF1 interactions (Figure 2E), revealing a competition between the two MEK isoforms for HSF1 binding and suggesting that MEK homodimers can interact with HSF1.

### ERK Suppresses MEK-HSF1 Interactions to Inactivate HSF1

To elucidate how ERK inactivates HSF1, we first examined the impact of ERK on MEK-mediated HSF1 activation. Whereas *ERK1* depletion promoted MEK-HSF1 interactions (Figure 2F), *ERK1* overexpression mitigated these interactions and suppressed HSF1 (Figures 2G and S2E). Thus, we contemplated three possible scenarios (Figure 2H): (1) both MEK substrates, ERK and HSF1, compete for MEK interaction; (2) ERK, like MEK, binds HSF1 and thereby competes for HSF1 interaction; and (3) ERK inhibits MEK kinase activity toward HSF1. Each of the first two scenarios predicts competition between two protein complexes; in contrast, the third scenario predicts that ERK assembles with MEK and HSF1 into a single protein complex. Interestingly, under HS HSF1 precipitated both MEK and ERK (Figure 2I), and ERK precipitated both MEK and HSF1 (Figure 2J). Although these results do not exclude the existence of independent MEK-ERK and MEK-HSF1 complexes, they argue against the two complexes being stable and prevalent, as depicted in the first scenario. To test the second scenario, we detected ERK-HSF1 interactions by PLA, as this scenario predicts HSF1 as a substrate for both ERK and MEK. The specificity of ERK antibodies was validated in *ERK1*-depleted cells (Figure S2F). In contrast to evident MEK1-ERK interactions (Figure 2K), no apparent PLA signals denoting ERK-HSF1 interactions were detected (Figure 2L), suggesting lack of direct contact between these two proteins. Moreover, while *ERK1* overexpression mitigated MEK-HSF1 interactions, less ERK1 proteins were precipitated with HSF1 (Figure 2G), conflicting with heightened ERK1-HSF1 interactions predicted by the second scenario. Thus, these results not only refute the second scenario but also suggest that ERK complexes with HSF1 via MEK, in line with the third scenario. Importantly, *MEK* depletion markedly diminished ERK-HSF1 coIP (Figure 2J). These results suggest existence of a protein complex comprising ERK, MEK, and HSF1, wherein ERK suppresses HSF1 indirectly, via inhibition of MEK.

### MEK Phosphorylates Ser326 to Activate HSF1

Under HS, HSF1 undergoes a series of phosphorylating events, among which Ser326 phosphorylation stimulates its activation (Guettouche et al., 2005). Yet, the identity of the kinase remains elusive. To determine whether MEK phosphorylates Ser326, we examined the effect of MEK blockade on this modification using a phosphospecific antibody that recognized HSF1<sup>WT</sup>, but not HSF1<sup>S326A</sup>, proteins (Figure S3A). Either *MEK* knockdown or U0126 treatment impaired Ser326 phosphorylation (Figures 3A and S3B). Conversely, a constitutively active mutant, MEK1<sup>DD</sup> (S218D/S222D) (Brunet et al., 1994a), induced Ser326 phosphorylation and activation of HSF1 without HS (Figures 3B and 3C). ERK inhibition enhanced Ser326 phosphorylation, and *MEK* depletion abolished this effect (Figures 3D and S3C), indicating MEK-dependent regulation. HSF1<sup>S326A</sup> mutants displayed impaired transcriptional activities (Figure 3E), congruent with their defective nuclear translocation and DNA-binding capacity (Figures S3D and 3F). Moreover, HSF1 proteins were reduced in *MEK*-deficient cells (Figures 3A and S3E). To determine whether MEK impacts HSF1 stability, we performed cycloheximide chase experiments. *MEK* depletion shortened

**Figure 2. ERK, MEK, and HSF1 Form a Stress-Inducible Protein Complex**

(A and B) After HS at 43°C for 30 min, endogenous HSF1 proteins were precipitated from HEK293T cells. WCL, whole cell lysate; HC, heavy chain.

(C) Endogenous MEK1-HSF1 interactions were detected by PLA in HeLa cells using a rabbit anti-MEK1 antibody and a mouse anti-HSF1 antibody. Scale bars represent 50 μm for LM, 10 μm for HM.

(D and E) Endogenous MEK-HSF1 interactions were detected by IP in HEK293T cells stably expressing shRNAs.

(legend continued on next page)

the half-life of HSF1<sup>WT</sup> protein; importantly, while HSF1<sup>S326A</sup> mutants were less stable in *MEK*-proficient cells, their stability was not evidently affected by *MEK* deficiency (Figure 3G), suggesting that *MEK* stabilizes HSF1 largely via Ser326 phosphorylation. These results indicate that *MEK* controls *in vivo* HSF1 Ser326 phosphorylation, a modification critical for its activation and stability.

In vitro, recombinant *MEK1* proteins directly phosphorylated HSF1 at Ser326, and U0126 blocked this event (Figure 3H). ERK was reported to phosphorylate HSF1 at Ser307 (Chu et al., 1998), implying a direct ERK-HSF1 interaction. To exclude direct Ser326 phosphorylation by ERK, we performed *in vitro* HSF1 phosphorylation using immunoprecipitated endogenous ERK complexes that would comprise ERK associated with or without *MEK*. Although precipitated complexes phosphorylated Ser326, this event was blocked by U0126, but not by FR180204 (Figure 3I). U0126, but not FR180204, blocked phosphorylation of recombinant ERK1 proteins by the same precipitates (Figure 3I). In contrast, FR180204, but not U0126, blocked phosphorylation of myelin basic protein (MBP), a known ERK substrate (Ahn et al., 1991), by the same precipitates (Figure 3I), showing ERK blockade by FR180204. These results strongly suggest that *MEK*, rather than ERK, directly phosphorylates Ser326. Moreover, recombinant ERK1 proteins impeded *in vitro* HSF1 Ser326 phosphorylation by recombinant *MEK1* proteins (Figure 3J), consistent with the suppression of Ser326 phosphorylation by ERK *in vivo*. In contrast, ERK promotes HSF1 Ser307 phosphorylation. ERK depletion diminished Ser307 phosphorylation, however, this effect was largely abolished in *MEK*-deficient cells (Figure S3F), again indicating *MEK* dependence. *MEK* inhibition impaired Ser326 phosphorylation but enhanced Ser307 phosphorylation of HSF1<sup>WT</sup> proteins (Figure S3G). Interestingly, phosphomimetic mutant HSF1<sup>S326D</sup> proteins displayed reduced basal Ser307 phosphorylation and resisted induction of this phosphorylation by *MEK* inhibition (Figure S3G). These results support that Ser326 phosphorylation by *MEK* represses Ser307 phosphorylation, and ERK impacts HSF1 Ser326 and Ser307 phosphorylation via *MEK* inhibition. Activated ERK phosphorylates Thr292/386 to inhibit *MEK1* (Brunet et al., 1994b). *MEK1*<sup>T292A,T386A</sup> mutants both heightened basal HSF1 Ser326 phosphorylation and blocked ERK-mediated suppression of this phosphorylation in cells depleted of endogenous *MEK* (Figure 3K), indicating that ERK suppresses Ser326 phosphorylation via feedback phosphorylation of *MEK*.

Interestingly, two *MEK1* mutations identified in human melanomas, P124S and E203K (Nikolaev et al., 2012), caused constitutive HSF1 phosphorylation and activation (Figures 3L and 3M). Conversely, in human melanoma cells, *MEK* inhibitors impaired constitutive HSF1 phosphorylation and binding to HSP promoters (Figures 3N and 3O). These results indicate that *MEK*

controls both inducible HSF1 activation in stressed cells and constitutive HSF1 activation in malignant cells.

### MEK Preserves Cellular Proteostasis

HSF1 could maintain cellular proteostasis via HSPs. To examine the impacts of HSF1 on protein folding, we employed the glucocorticoid receptor (GR) as a model. Proper GR folding depends on HSP90 and misfolded proteins are cleared by the ubiquitin-proteasome system (Taipale et al., 2010). HSF1 knockdown induced GR-GFP ubiquitination and depletion (Figure S4A), indicating protein destabilization by HSF1 deficiency. This resulted from diminished cellular chaperoning capacity, as lysates of HSF1-deficient cells were less efficient in reactivating denatured luciferase (Figure S4B). Similarly to HSF1 deficiency, *MEK* blockade depleted GR-GFP, and this depletion is not due to GFP instability or general expression changes, since co-expressed GFP was not affected (Figure 4A). Instead, *MEK* blockade ubiquitinated GR-GFP (Figure 4B). This is not due to impaired proteasomal function, as proteasome inhibition by Bortezomib caused GR-GFP accumulation and *MEK* inhibitors did not affect proteasomal activities (Figures S4C–S4E). In fact, AZD6244 and *MEK* knockdown both depleted chaperoning capacity (Figures 4C and S4F), revealing modulation of protein folding and stability by *MEK*.

In line with a key role of HSF1 in governing cellular proteome, HSF1 depletion induced protein Lys48-specific ubiquitination, a modification marking proteins for proteasomal degradation (Pickart and Eddins, 2004), in both detergent-soluble and -insoluble fractions (Figure S4G). This change suggests global protein destabilization. Consistent with HSF1 inactivation, AZD6244 diminished Ser326 phosphorylation, reduced HSPs, and induced overall ubiquitination (Figure 4D). Overnight AZD6244 treatment also destabilized HSF1 (Figures 4D and S4H). *MEK* knockdown induced global ubiquitination as well (Figure S4I). Importantly, AZD6244 failed to deplete HSF1 and provoke ubiquitination in cells stably overexpressing HSF1<sup>S326D</sup> (Figures 4E and S4J), indicating a causative role of HSF1 inactivation in protein instability due to *MEK* inhibition. *In vivo* *MEK* inhibition also depleted HSPs and HSF1 and provoked ubiquitination in primary tissues (Figure 4F).

To investigate ubiquitomic changes due to *MEK* inhibition, we conducted mass spectrometry (MS)-based analyses of ubiquitinated peptides enriched by a novel ubiquitin branch motif antibody (Figure 4G) (Kim et al., 2011). We compared the ubiquitomes of A2058 cells treated with and without AZD6244 for 8 hr. In total, 3,425 non-redundant ubiquitinated peptides, assigned to 1,715 distinctive proteins, were profiled (Figures 4H and S4K; Table S1). AZD6244 both increased and decreased peptide ubiquitination (Figure 4H). When a 2.5-fold cutoff was defined as the significant change, a collection of 76 non-redundant

(F) Endogenous MEK-HSF1 interactions were detected in HEK293T cells transfected with siRNAs.

(G) Endogenous HSF1-MEK and HSF1-GFP-ERK1 interactions were detected in HEK293T cells transfected with indicated plasmids.

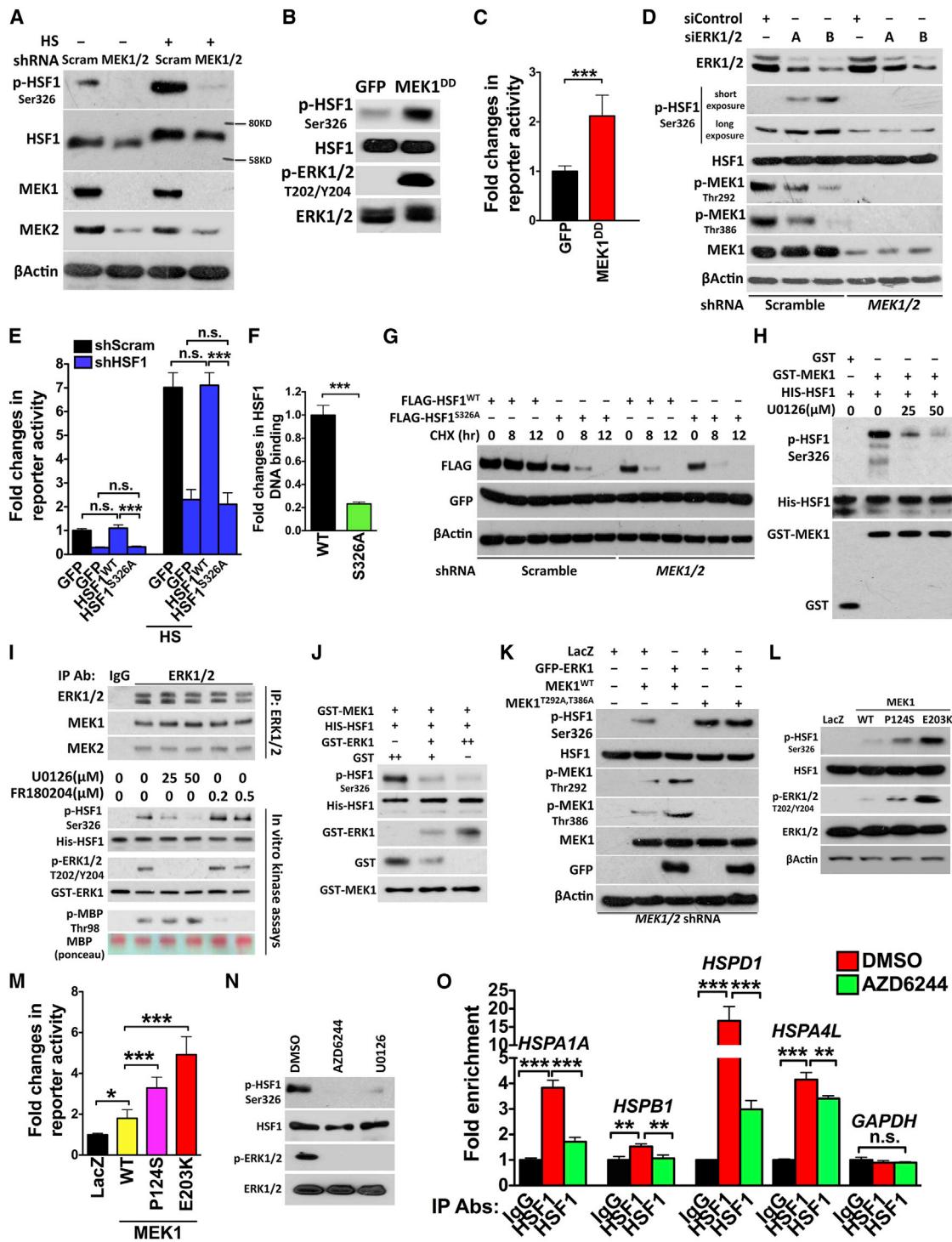
(H) Schematic depiction of three possible scenarios. P, phosphorylation.

(I) Immediately after HS, HSF1-ERK interactions were detected by coIP.

(J) Endogenous ERK-HSF1 interactions were detected in HEK293T cells stably expressing shRNAs.

(K and L) Endogenous ERK-MEK and ERK-HSF1 interactions were detected by PLA in HeLa cells. Scale bars represent 50 μm for LM, 10 μm HM.

See also Figure S2.

**Figure 3. MEK Phosphorylates Ser326 to Activate HSF1**

(A and B) HSF1 Ser326 phosphorylation was measured by immunoblotting in HEK293T cells stably expressing shRNAs or transfected with MEK1<sup>DD</sup> plasmid.

(C) GFP or MEK1<sup>DD</sup> plasmids were co-transfected with dual HSF1 reporter plasmids into HEK293T cells (mean  $\pm$  SD, n = 5, Student's t test).

(D) Control or ERK-targeting siRNAs, A (siERK1\_1 and siERK2\_1) and B (siERK1\_3 and siERK2\_2), were transfected into HEK293T cells stably expressing shRNAs.

(E) GFP or FLAG-HSF1 plasmids were co-transfected with dual HSF1 reporter plasmids into HEK293T cells stably expressing shRNAs (mean  $\pm$  SD, n = 6, ANOVA).

(legend continued on next page)

peptides were distinguished. These peptides represent 68 unique proteins that perform diverse molecular functions and engage in a wide array of biological processes (Figures 4I and 4J). Intriguingly, the most enriched pathway was translation elongation (Figure 4J). Analyses of these 68 proteins revealed a functional association network encompassing three subnetworks (Figure 4K). Of particular interest is the “translation” subnetwork that consists of 7 ribosomal subunit proteins, highlighting a prominent impact of MEK on ribosome machinery. Moreover, embedded within the network are several oncogenes and tumor suppressors, including c-MYC, Cyclin D1, HIF1A, TP53, and NF1 (Figure 4K). Although the mechanisms whereby MEK regulates these key players in oncogenesis are likely multifaceted, accumulating evidence has implicated HSPs in modulating their stabilities (Isaacs et al., 2002; Müller et al., 2004). Thus, MEK could impact these proteins at least in part via HSF1.

To validate our MS findings, we elected several target proteins. Torsin-1A interacting protein 2 (TOR1AIP2) and ribosomal protein L3 (RPL3) exhibited 61.0- and 13.7-fold increases, respectively, in ubiquitination (Table S1). To facilitate detection, we expressed V5-tagged TOR1AIP2 and RPL3 proteins via a constitutive promoter. AZD6244 treatment for 8 hr did not alter levels of both V5-tagged proteins but increased their ubiquitination (Figures 4L and 4M). Our MS results also revealed decreased ubiquitination of proteins including c-MYC, RPL15, RPL24, and RPS20 (Table S1). We confirmed reductions in both ubiquitination and total levels of endogenous c-MYC proteins (Figure 4N). Similar results were also observed for V5-tagged RPL15, RPL24, and RPS20 (Figures S4L–S4N), suggesting shortened protein half-life. Indeed, proteasome blockade by MG132 prevented RPL15-V5 depletion by AZD6244 and revealed its elevated ubiquitination (Figure 4O). Consistent with a critical role of HSF1 inactivation, both V5-tagged RPL15 and RPL3 proteins were highly ubiquitinated following HSF1 knockdown, and AZD6244 subtly affected this ubiquitination (Figures 4O and 4P). Importantly, MG132 prevented depletions of endogenous RPL15 and RPL3 by AZD6244 and MEK knockdown (Figures 4Q and S4O), confirming destabilization of ribosomal proteins by MEK deficiency. While HSF1 knockdown diminished endogenous RPL15 and RPL3, HSF1<sup>S326D</sup> expression elevated their basal levels and protected them from AZD6244-induced depletions (Figures 4R and S4P). These findings together indicate that MEK inhibition inactivates HSF1 to deplete cellular

chaperoning capacity. In consequence, protein destabilization and ubiquitomic imbalance ensue.

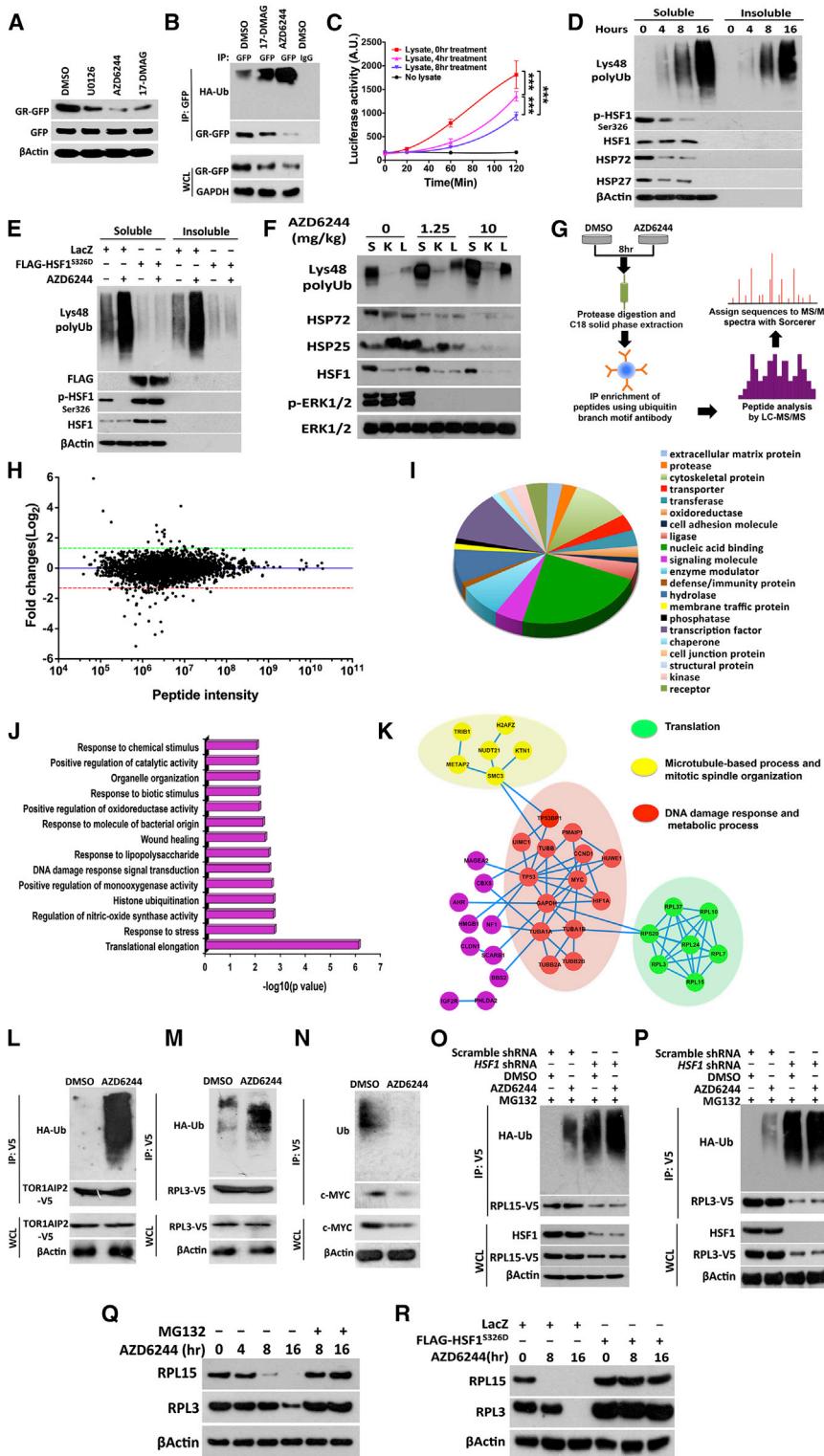
### MEK Inhibition Provokes Aggregation and Amyloidogenesis in Malignant Cells

Increased ubiquitination in detergent-insoluble fractions suggests protein aggregation (Figure 4D). To demonstrate this, we performed ubiquitin immunostaining. Melanoma cells were treated with Bortezomib and AZD6244 for 24 hr to provoke evident aggregation. As expected, bright fluorescent punctate foci emerged in Bortezomib-treated cells (Figure 5A), demarcating ubiquitin-containing aggregates. AZD6244 depleted HSF1 and, albeit to a lesser extent, induced punctate foci (Figures 5A and S5A). We theorized that blockade of proteasomal degradation of AZD6244-induced misfolded proteins would aggravate aggregation. Indeed, Bortezomib co-treatment augmented AZD6244-induced punctate foci (Figure 5A). We further confirmed the impact of MEK on aggregation using an expanded polyglutamine tract protein (polyQ79) (Sánchez et al., 2003). Proteins with expanded polyQ fragments are aggregation-prone and causally related to neurodegenerative disorders (Orr and Zoghbi, 2007). PolyQ79-expressing cells, as expected, contained large aggregates (Figure S5B). HSF1 knockdown and AZD6244 both enlarged polyQ aggregates; importantly, HSF1<sup>S326D</sup> expression antagonized AZD6244-induced aggregate enlargement (Figures 5B and S5C). Bortezomib also enhanced aggregation, and combined treatment produced the largest aggregates (Figure 5B). Thus, both MEK and the proteasome suppress protein aggregation.

Aggregation-prone proteins can form amyloid fibrils (AFs) enriched for  $\beta$  sheet structures (Eisenberg and Jucker, 2012). To assess whether HSF1 and MEK impact amyloid formation, we stained polyQ79-expressing cells with Thioflavin T (ThT) and Congo red (CR), two fluorescent dyes widely used to diagnose amyloids (Chiti and Dobson, 2006). PolyQ79 expression enhanced ThT and CR staining (Figures S5D and S5E), as expected. AZD6244, Bortezomib, and combined treatment further intensified this staining, and HSF1<sup>S326D</sup> expression antagonized the effect of AZD6244 (Figure 5C). Treatments also enhanced ThT and CR staining of human tumor cell lines (Figures 5D and S5F), suggesting emergence of endogenous amyloid-like structures. The presence of soluble amyloid oligomers (AOs) confirmed amyloidogenesis. AOs are believed to constitute a

- (F) FLAG-HSF1 plasmids were transfected into HEK293T cells stably expressing HSF1-targeting shRNAs. HSF1-DNA binding was measured after HS as described in Figure 1E using anti-FLAG antibodies. The results were normalized against nuclear FLAG-HSF1 levels (mean  $\pm$  SD, n = 3, Student's t test).
- (G) FLAG-HSF1 proteins were detected in HEK293T cells treated with 20  $\mu$ g/ml cycloheximide (CHX). Co-expressed GFP proteins served as internal controls.
- (H) Purified GST-MEK1 proteins (100 ng) were incubated with U0126 at RT for 20 min followed by incubation with 400 ng purified His-HSF1 proteins at RT for 30 min. HSF1 phosphorylation was detected by immunoblotting.
- (I) ERK complexes precipitated from HEK293T cells were treated with U0126 or FR180204, followed by incubation with 400 ng His-HSF1, 400 ng GST-ERK1, or 1,000 ng MBP proteins.
- (J) Inactive GST-ERK1 proteins were incubated with 100 ng GST-MEK1 and 400 ng His-HSF1 proteins at RT for 30 min.
- (K) LacZ or GFP-ERK1 plasmid was co-transfected with MEK1<sup>WT</sup> or MEK1<sup>T292A,T386A</sup> plasmid into HEK293T cells stably expressing MEK-targeting shRNAs.
- (L) HSF1 Ser326 phosphorylation was detected in HEK293T cells transfected with indicated plasmids.
- (M) HSF1 activities were measured by the dual reporter system in HEK293T cells transfected with indicated plasmids (mean  $\pm$  SD, n = 6, ANOVA).
- (N) WM115 cells were treated with 20 nM AZD6244 or 20  $\mu$ M U0126 overnight.
- (O) HSF1 chromatin immunoprecipitation (ChIP) assays were performed using WM115 cells treated with DMSO or 20 nM AZD6244 overnight. The results were normalized against the values of IgG controls (mean  $\pm$  SD, n = 3, ANOVA).

See also Figure S3.

**Figure 4. MEK Preserves Proteostasis**

(A) GFP and GR-GFP plasmids were co-transfected into HEK293T cells followed by treatments with 20 nM AZD6244, 20 μM U0126, or 200 nM 17-DMAG for 4 hr.

(B) GR-GFP plasmids were co-transfected into HEK293T cells with HA-Ub-K48 plasmids, which encode a mutant ubiquitin that can be conjugated to protein substrates only via lysine 48. Following treatments with 20 nM AZD6244 or 200 nM 17-DMAG for 4 hr, GR-GFP proteins were precipitated and ubiquitination was detected using anti-HA antibodies.

(C) Denatured firefly luciferases were incubated with lysates of A2058 cells treated with DMSO or 20 nM AZD6244 (mean ± SD, n = 4, ANOVA).

(D) A2058 cells were treated with 20 nM AZD6244, and ubiquitinated proteins were detected in both detergent-soluble and detergent-insoluble fractions using Lys48-specific ubiquitin antibodies.

(E) A2058 cells stably expressing LacZ or HSF1S326D were treated with 20 nM AZD6244 for 8 hr.

(F) C57BL/6J mice were i.p. injected with DMSO or AZD6244 three times a week for 2 weeks. S, spleen; K, kidney; L, liver.

(G) Experimental procedures of MS-based quantitation of ubiquitinated peptides, two technical replicates per treatment.

(H) Scatter plot of relative changes in peptide abundance between treated and control conditions. The green and red lines indicate 2.5-fold cutoffs.

(I) The classification of the 68 proteins was performed using the PANTHER gene list analysis tool (<http://www.pantherdb.org>).

(J) The Gene Ontology (GO) biological process enrichment analysis was performed using the web-based Enrichr software application.

(K) Interaction network of the 68 proteins. Known and predicted protein interactions were derived from the STRING database (<http://www.string-db.org>), and the network was visualized using Cytoscape software.

(L and M) TOR1AIP2-V5 or RPL3-V5 plasmids were co-transfected with HA-Ub-K48 plasmids into HEK293T cells. Following 20 nM AZD6244 treatment for 8 hr, proteins were precipitated with anti-V5 antibodies.

(N) Following AZD6244 treatment, endogenous c-MYC proteins were precipitated from A2058 cells and immunoblotted with anti-ubiquitin antibodies. (O and P) RPL15-V5 and RPL3-V5 plasmids were co-transfected with HA-Ub-K48 plasmids into HEK293T cells stably expressing shRNAs. Cells were treated with 500 nM MG132 alone or co-treated with 20 nM AZD6244 for 8 hr.

(Q) Endogenous RPL15 and RPL3 proteins were detected in A2058 cells treated with 20 nM AZD6244 alone or co-treated with 500 nM MG132. (R) Endogenous RPL15 and RPL3 proteins were detected in A2058 cells stably expressing LacZ or HSF1S326D with AZD6244 treatment.

See also Figure S4 and Table S1.

key toxic species in neurodegenerative disorders and can be detected by the conformation-dependent antibody A11 (Chiti and Dobson, 2006; Glabe, 2008; Kayed et al., 2003). Treatments

not only exaggerated AO induction by polyQ79, but also provoked genesis of endogenous AOs in human tumor cell lines (Figures 5E and 5F). Similarly to AZD6244, HSF1 depletion

induced endogenous AOs and AFs (Figures S5G and S5H). A previously characterized antibody, OC, was used to detect AFs (Kayed et al., 2007). Again, HSF1<sup>S326D</sup> expression suppressed AZD6244-induced amyloidogenesis (Figures 5G and 5H).

A unique feature of amyloids is their ability to seed AFs (Chiti and Dobson, 2006). In amyloid seeding experiments, lysates of HSF1-depleted cells accelerated formation of A $\beta$  AFs (Figure S5I). Similarly, lysates of cells treated with AZD6244, Bortezomib, and combination all exhibited augmented seeding efficacy (Figure 5I), which was confirmed using OC antibodies (Figure S5J). HSF1<sup>S326D</sup> expression abolished the effect of AZD6244 (Figure 5I). Furthermore, transmission electron microscopy revealed that while seeding with DMSO-treated cell lysates resulted into disordered rod-like protofibrils, a dense lattice of fiber-like structures emerged following seeding with AZD6244-treated lysates; in comparison, seeding with Bortezomib-treated lysates produced mature fibrils, and a compacted network of fibrils were assembled after seeding with lysates from combined treatment (Figure 5J).

The amyloidogenic effects of AZD6244 and Bortezomib were validated genetically. Depletion of the  $\beta$ 5 subunit (PSMB5) of the 26S proteasome, a primary target of Bortezomib (Oerlemans et al., 2008), caused accumulation of ubiquitinated proteins (Figure S5K). Mimicking pharmacological inhibitors, genetic depletions of MEK, PSMB5, or both all provoked amyloidogenesis (Figures 5K and S5L–S5Q).

To determine whether amyloids contribute to inhibitor-induced toxicities, we blocked amyloidogenesis with ThT, which impedes amyloid fibrillization via physical binding (Alavez et al., 2011). In melanoma cells, ThT suppressed amyloid induction by inhibitors and improved cellular growth and survival by 50% (Figures 5L–5N). CR treatment and neutralization of AOs with A11 antibodies exerted similar protection (Figures 5O and S5R). Congruent with mitigated amyloidogenesis, HSF1<sup>S326D</sup> expression not only stimulated the growth of melanoma cells but also rendered them refractory to MEK inhibition (Figure 5P).

Surprisingly, AZD6244 did not induce AOs in primary mouse embryonic fibroblasts (MEFs) and tissues (Figures 5Q and S5S). This is not due to inability to detect murine amyloids, as severe stress did induce AOs in murine cells (Figure S5T). These results suggest that non-transformed cells may be more refractory to amyloidogenesis than malignant cells. To assess this, we compared AO levels in primary human mammary epithelial cells (PHMC), immortalized human mammary epithelial (MCF10A) cells, and tumorigenic mammary epithelial (MCF7) cells treated with AZD6244, Bortezomib, and the combination. Each of these three treatments caused marked AO induction in MCF7 cells, slight induction in MCF10A cells, and no induction in PHMC (Figure 5R). A similar pattern was observed in primary human Schwann cells (PHSC) and their malignant counterparts, 90-8TL and S462 cells (Figure 5S). Immortalized and transformed cells, except S462, also showed elevated basal levels of AOs (Figures 5R and 5S). Intriguingly, AO levels positively correlated with malignant states (Figure 5T), supporting proteomic imbalance as an intrinsic feature of malignancy. We theorized that the lack of elevated basal AOs in S462 cells might be due to amyloid-associated toxicity. Indeed, blockade of cell death by a pan-caspase inhibitor (CI) elevated AO levels in immortalized and transformed cells, revealing height-

ened amyloidogenesis; in contrast, it did not elevate AOs in primary cells (Figures 5U and 5V), supporting an absence of amyloidogenesis. These results indicate that malignant cells are distinctively susceptible to amyloidogenesis.

### Combined Proteasome and MEK Inhibition Disrupts Tumor Proteostasis and Suppresses Malignancy

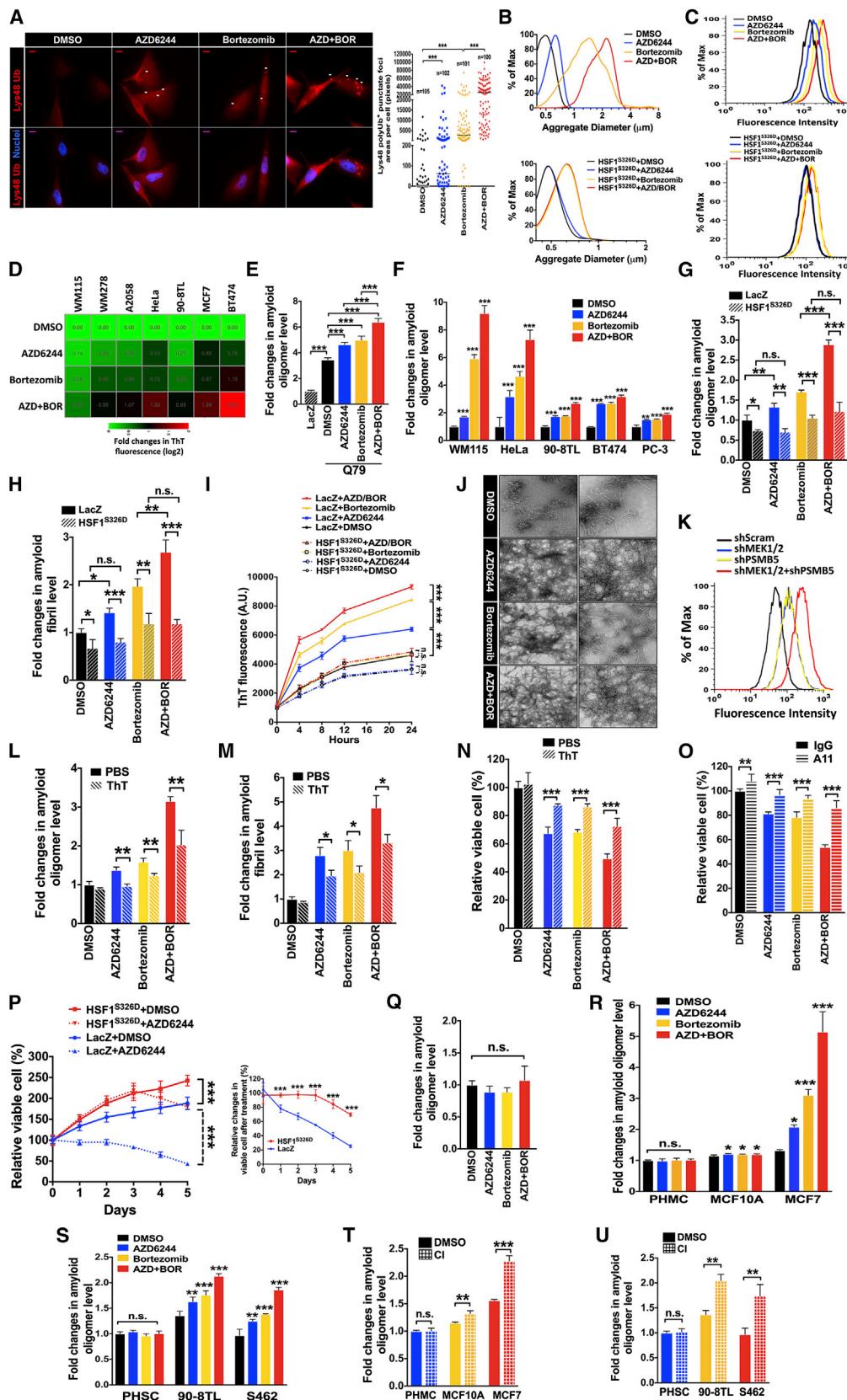
MEK and proteasome inhibition, individually, disturbed proteostasis in tumor cells to certain degrees, however, the combination of both augmented this effect and, accordingly, markedly impaired the growth and survival of human tumor cell lines (Figure 6A). Of note, this combination did not impact primary cells (Figure 6A).

In vivo, whereas low doses of AZD6244 or Bortezomib alone exhibited no significant impacts on xenografted melanomas, the combination potently retarded their growth (Figures 6B and S6A). All mice receiving the combined treatment remained alive and their body weights remained constant; in contrast, all mice in the other groups died and lost ~25% of body weight (Figures 6C and S6B). AZD6244 or Bortezomib alone slightly elevated ubiquitination in tumors, however, the combination markedly aggravated this effect (Figures 6D and S6C). While Bortezomib induced HSF1 Ser326 phosphorylation and HSP expression, AZD6244 co-treatment suppressed this stress response and induced caspase 3 cleavage (Figures 6D and S6C). Accordingly, AOs were evidently elevated in tumors receiving combined treatment (Figure 6E). Of particular interest is an inverse correlation between amounts of AOs and tumor masses (Figure S6D), supporting an adverse impact of AOs on malignant growth. Congruent with amyloidogenesis, tumors receiving combined treatment displayed potent seeding capacities and enhancement of CR staining (Figures 6F and 6G). Intratumoral AFs were further demonstrated by the hallmark birefringence of CR staining (Figure 6H), ThT labeling (Figure S6E), and immunostaining with OC antibodies (Figure 6I). In contrast, the combined treatment did not induce AOs and apoptosis in primary tissues of the same tumor-bearing mice, despite elevated ubiquitination (Figures S6F and S6G).

To investigate whether the combined treatment impedes experimental metastasis, we intravenously injected melanoma cells expressing a luciferase transgene into NOD/SCID mice. During a 6-week period, only mice receiving combined treatment gained body weight (Figure 6J), suggesting improved health. In vivo imaging detected luminescence in 40% of mice treated with DMSO, AZD6244, or Bortezomib alone (Figure 6K). Histological examination confirmed melanoma metastases to the lung, skeletal muscle, adipose tissue, and ovary (Figure 6L; Table S2). In contrast, none of the mice receiving combined treatment displayed discernible luminescence or metastases (Figures 6K and 6M). Together, these results demonstrate that combined MEK and proteasome inhibition provokes proteotoxic stress and amyloidogenesis within tumors and exerts robust anti-neoplastic effects.

### Amyloidogenesis Is Tumor Suppressive

Evident apoptosis in tumor regions showing intense CR staining suggests a causative role of amyloidogenesis in treatment-induced toxicity (Figure 7A). To confirm this, we blocked



(legend on next page)

intratumoral amyloid induction via in vivo CR administration. CR not only accelerated melanoma growth but also potently antagonized the tumor suppression imposed by combined MEK and proteasome inhibition (Figure 7B). Penetration of CR into tumor tissues was indicated by intense light absorption of tumor lysates at 498 nm (Figure 7C), a characteristic of this amyloid stain (Sladewski et al., 2006). Congruent with enhanced malignancy, CR treatment enlarged tumor masses, deteriorated body conditions, and shortened animal survival (Figures 7D, S7A, and S7B).

While CR reduced amyloids in tumor tissues, it did not diminish ubiquitination (Figures 7E–7G). These results indicate no interference of CR with MEK and proteasome inhibitors and further support a specific action of CR in blocking amyloid genesis. In accordance with accelerated growth, CR-treated tumors displayed reduced caspase 3 cleavage (Figure 7G). Collectively, these results strongly suggest that amyloidogenesis is tumor suppressive and evidently contributes to the anti-neoplastic effects of combined MEK and proteasome inhibition.

## DISCUSSION

### HSF1 Is a New MEK Substrate

Unexpectedly, our results reveal HSF1 as a physiological substrate for MEK, challenging the prevailing paradigm wherein ERK exclusively instigates the effects of RAS-RAF-MEK signaling. Our results further show that MEK activates but ERK inactivates HSF1. Importantly, our findings integrate these two seemingly contradictory actions and support the assembly of a ternary ERK-MEK-HSF1 protein complex. In aggregate, our findings propose a bifurcated, rather than a linear, RAS-RAF-MEK cascade. MEK, as a central nexus, both conveys upstream stimuli and governs two discrete but interconnected downstream effector pathways, of which one is mediated by ERK and the other by HSF1 (Figure 7H). In a negative feedback fashion,

ERK finely attunes HSF1 activation via inhibitory phosphorylation of MEK (Figure 7H). While our studies focused on MEK-mediated Ser326 phosphorylation, other kinases can also regulate HSF1.

### Guarding of Proteostasis by RAS-RAF-MEK Signaling

Our findings uncover a new function of RAS-RAF-MEK signaling in regulating proteostasis. Diverse proteotoxic stressors commonly activate MEK (Figure 1A). Through HSF1 activation, RAS-RAF-MEK signaling heightens cellular chaperoning capacity to guard proteomic integrity.

MEK-HSF1 regulation could have key physiological implications. Mitogens stimulate RAS/MAPK signaling and downstream mTORC1 (Laplante and Sabatini, 2012). However, heightened protein synthesis driven by mTORC1 encumbers cellular protein quality-control machinery. It thus appears necessary for mitogens, via MEK, to concurrently mobilize the HSF1-controlled chaperone system to ensure productive protein synthesis and, thereby, avert proteomic imbalance. Interestingly, MEK also governs translation capacity via HSF1 (Figure 4K). Thus, RAS-RAF-MEK signaling synchronizes protein quantity- and quality-control machineries to support cellular growth.

It is also tempting to speculate that RAS-RAF-MEK signaling may antagonize protein-misfolding diseases, such as amyloidosis, via guarding proteostasis.

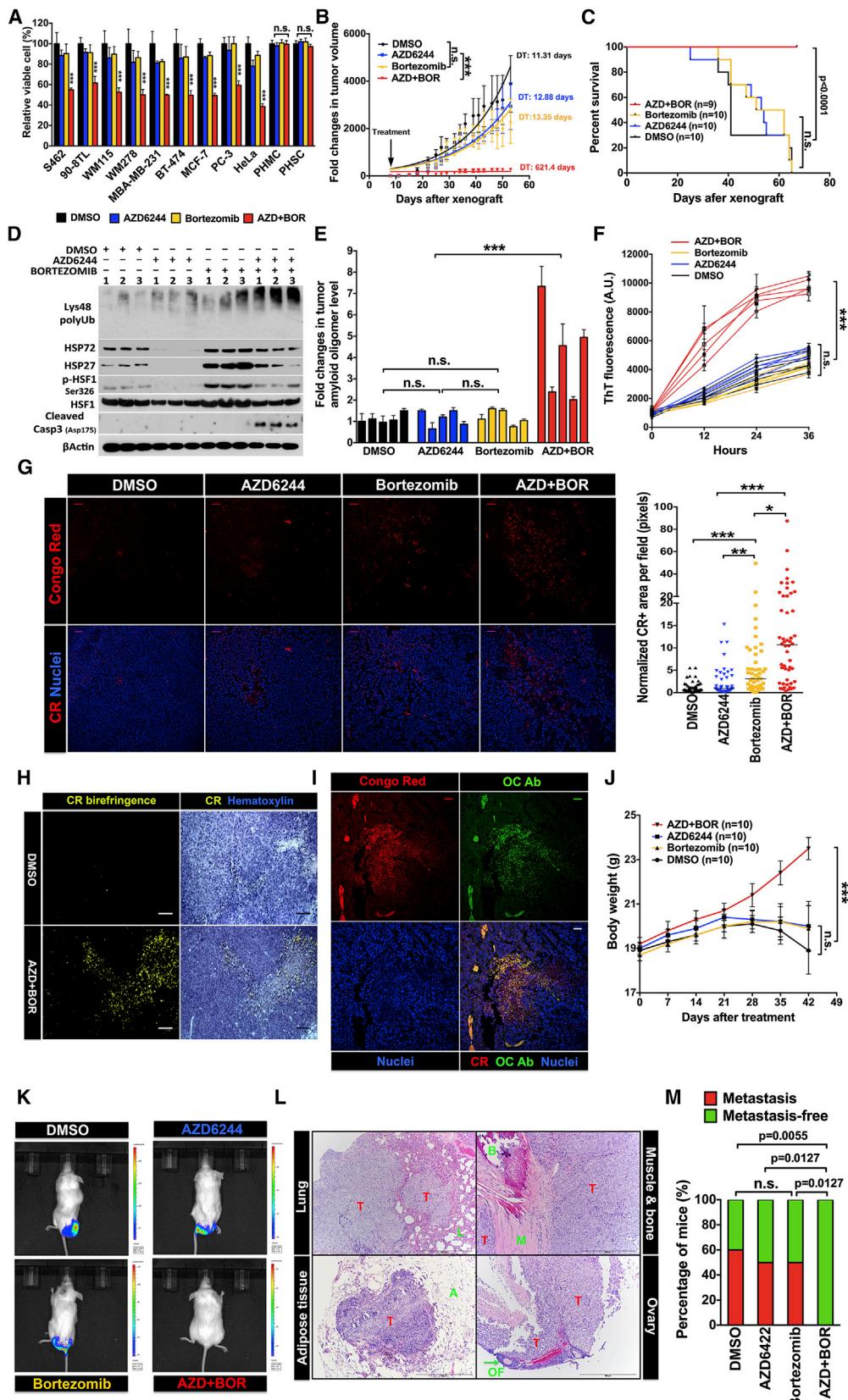
### Proteomic Instability of Cancer

Our findings pinpoint a pro-amyloidogenic nature of malignant state. The susceptibility of malignant cells to amyloid genesis likely originates from their debilitated proteostatic state, which is particularly vulnerable to perturbations. Unlike non-transformed cells, malignant cells constantly endure proteomic imbalance, evidenced by elevated basal levels of amyloids (Figures 5T and 5U). Accordingly, HSF1, otherwise latent in primary cells, is constitutively mobilized in tumor cells to retain the fragile

### Figure 5. MEK and Proteasome Inhibition Provoke Protein Aggregation and Amyloidogenesis

- (A) WM115 cells treated with 20 nM AZD6244, 100 nM Bortezomib, or both for 24 hr were stained with Lys48-specific ubiquitin antibodies. Arrowheads mark ubiquitin-positive aggregates. Scale bar represents 10 μm. Amounts of aggregates per cell were quantitated using ImageJ (median, n ≥ 100, ANOVA).
- (B and C) Following transfection with polyQ79 plasmids alone or with both polyQ79 and HSF1<sup>S326D</sup> plasmids for one day, HEK293T cells were treated with inhibitors as described in (A). Cells were either analyzed for aggregate size or stained with 10 μM ThT.
- (D) Treated tumor cell lines were stained with 10 μM ThT. Geometric means were used to calculate fold changes in ThT fluorescence intensity and the log<sub>2</sub>(FC) values were presented as a heat map.
- (E) HEK293T cells were transfected with LacZ or polyQ79 plasmid. Following treatments, AOs were quantitated by ELISA using A11 antibodies (mean ± SD, n = 3, ANOVA).
- (F) Intrinsic AOs were detected in human tumor cell lines (mean ± SD, n = 3, ANOVA).
- (G and H) A2058 cells stably expressing LacZ or HSF1<sup>S326D</sup> were treated for 24 hr. Amyloids were quantitated by ELISA (mean ± SD, n = 3, Student's t test).
- (I) Synthetic Aβ1-42 peptides (20 μM) were incubated at RT with gentle shaking with 20 μg lysates of A2058 cells treated with inhibitors. AF formation was monitored by ThT binding (mean ± SD, n = 3, ANOVA).
- (J) For TEM studies (left panel, 80,000×; right panel, 200,000×), 20 μM synthetic Aβ1-42 peptides were incubated with A2058 cell lysates in PBS at 37°C with gentle shaking for 2 days. Scale bars represent 100 nm.
- (K) HEK293T cells stably expressing different shRNAs were stained with 10 μM ThT.
- (L and M) After pre-incubation with 10 μM ThT for 6 hr, A2058 cells were treated for 24 hr. Amyloids were quantitated (mean ± SD, n = 3, Student's t test).
- (N and O) After pre-incubation with 10 μM ThT for 6 hr or transfection with 100 ng A11 antibodies using JBS-Proteoducin for 16 hr, A2058 cells were treated for 24 hr. Viable cells were quantitated using CellTiter-Blue reagents (mean ± SD, n = 6, Student's t test).
- (P) A2058 cells stably expressing LacZ or HSF1<sup>S326D</sup> were treated with DMSO or 20 nM AZD6244. Viable cells were quantitated (mean ± SD, n = 6, ANOVA). Relative changes in viable cells after treatment were calculated by normalizing the values of AZD6244-treated cells against the values of DMSO-treated cells at each time point.
- (Q–S) Following treatments with 20 nM AZD6244, 100 nM Bortezomib, or both for 24 hr, AOs were quantitated in primary MEFs and human cells (mean ± SD, n = 3, ANOVA).
- (T and U) Cells were treated with 50 μM Q-VD-OPh overnight and AOs were quantitated (mean ± SD, n = 3, Student's t test).

See also Figure S5.



proteomic equilibrium (Figure 7H). Proteomic chaos inevitably ensues following either HSF1 inactivation or even mild proteotoxic insults. Hence, tumors cells rely on HSF1 to sustain their malignant phenotypes (Dai et al., 2007). In contrast, the lack of intrinsic proteotoxic stress empowers primary cells to effectively buffer intense proteomic fluctuations and thereby avert deleterious consequences—aggregation and amyloidogenesis.

Excitingly, the distinct susceptibilities to proteomic perturbation between primary and malignant cells may be exploited to combat malignancy. Our findings support important roles for proteotoxic stress and amyloidogenesis in the toxicity of MEK inhibition in malignancy. Through protein destabilization, MEK inhibitors act as a proteotoxic stressor, mechanistically distinct from proteasome inhibitors. When applied as single agent, a MEK or proteasome inhibitor is incompetent to distress tumor proteostasis. However, combinatorial application exerts a profound impact, eliciting amyloidogenesis. Importantly, our findings strongly suggest a tumor-suppressive nature of amyloidogenesis (Figures 5L–O and 7B). These findings imply that amyloidogenesis, indicative of grave proteomic imbalance, may be of prognostic value in monitoring tumor progression and evaluating therapeutic responses. Conceptually, our findings suggest that proteomic instability is an intrinsic characteristic associated with malignant state and that, therefore, disruption of fragile tumor proteostasis may be a feasible therapeutic strategy.

## EXPERIMENTAL PROCEDURES

### Proximity Ligation Assay

Cells were fixed with 4% formaldehyde in PBS for 15 min at room temperature (RT). After blocking with 5% goat serum in PBS with 0.3% Triton X-100, cells were incubated with a pair of rabbit and mouse primary antibodies 1:200 diluted in the blocking buffer overnight at 4°C. Following incubation with Duolink PLA anti-rabbit Plus and anti-mouse Minus probes (OLINK Bioscience) at 37°C for 1 hr, ligation, rolling circle amplification, and detection were performed using Duolink In Situ Detection Reagents Red (OLINK Bioscience). Nuclei were stained with Hoechst 33342. Signals were visualized using a Leica TCS SP5 confocal microscope.

### Figure 6. Combined MEK and Proteasome Inhibition Exerts Potent Tumor-Suppressive Effects

- (A) After treatments with 20 nM AZD6244, 100 nM Bortezomib, or both for 24 hr, viable cells were quantitated using CellTiter-Blue reagents (mean ± SD, n = 6, ANOVA).
- (B and C) A2058 cells ( $1 \times 10^6$ ) were s.c. injected into NOD/SCID mice. After 7 days, mice were treated with DMSO, 5 mg/Kg AZD6244, 0.5 mg/Kg Bortezomib, or the combination via i.p. injection three times a week. Tumor volumes were measured using a caliper weekly (mean ± SEM, ANOVA). Tumor growth curves were fitted to exponential growth models to derive tumor-doubling time (DT). Kaplan-Meier survival curve was plotted for each group (Log-rank test).
- (D) Proteins were detected by immunoblotting, three tumors per group.
- (E) Tumor lysates were used to quantitate AOs, five tumors per group (mean ± SD, n = 3, ANOVA).
- (F) Tumor lysates were used to seed Aβ1–42 peptides, five tumors per group (mean ± SD, n = 3, ANOVA).
- (G) Tumor sections were stained with CR, five tumors per group. Ten random fields were taken for each section. Scale bar represents 50 μm. Total CR fluorescence in each field was quantitated using ImageJ and normalized against total nuclei (median, n = 50, ANOVA).
- (H) Following CR staining, tumor sections were visualized under polarized light microscopy. Scale bar represents 50 μm.
- (I) Following staining with AF-specific antibodies (OC), sections of tumors receiving combined treatment were further stained with CR. Scale bar represents 50 μm.
- (J) A2058 cells ( $1 \times 10^6$ ) stably expressing firefly luciferase transgene were intravenously (i.v.) injected into NOD/SCID mice. Treatments were initiated 1 day after as described in (B) for 6 weeks. Body weights were monitored weekly (mean ± SD, n = 10, ANOVA).
- (K) Detection of metastases by *in vivo* bioluminescence imaging.
- (L) Representative micrographs illustrate metastatic melanomas in the lung, skeletal muscle, pelvic adipose tissue, and ovary. T, tumors; L, lung; M, muscle; B, bone; A, adipose tissue; OF, ovarian follicle. Scale bar represents 500 μm.
- (M) Combined MEK and proteasome inhibition prevents experimental melanoma metastasis (Barnard's exact test).

See also Figure S2 and Table S2.

### CR and ThT Staining of Tumor Sections

Following deparaffinization and rehydration, tumor sections were stained with 0.5% CR in PBS at RT for 20 min followed by differentiation in alkaline solutions (0.01% NaOH, 50% alcohol). Nuclei were stained with either Hoechst 33342 or hematoxylin. Fluorescence was visualized using a Leica TCS SP5 confocal microscope and the birefringence visualized using a Leica DM5000B upright microscope equipped with polarized light filters. For ThT staining, sections were stained with 0.2% ThT in PBS at RT for 10 min, rinsed in 1% acetic acid for 2 min, and washed with ddH<sub>2</sub>O three times. Nuclei were stained with SYTO 62 (Life Technologies).

### Melanoma Xenograft Models

A2058 cells were subcutaneously (s.c.) injected into the left flanks of 9-week-old female NOD.CB17-Prkdc<scid>/J (NOD/SCID) mice (The Jackson Laboratory). For CR treatment, mice were intraperitoneally (i.p.) injected with PBS or CR 1 day prior to combined AZD6244 and Bortezomib treatments. Tumor volumes were calculated following the formula  $4/3\pi R^3$ . For experimental metastasis, engineered A2058 cells were transplanted into 10-week-old female NOD/SCID mice via tail vein injections. All mouse experiments were performed under a protocol approved by The Jackson Laboratory Animal Care and Use Committee.

### Statistical Methods

All statistical analyses were performed using Prism 6.0 (GraphPad software). Statistical significance: \*p < 0.05; \*\*p < 0.01; \*\*\*p < 0.001.

See also the Extended Experimental Procedures.

## SUPPLEMENTAL INFORMATION

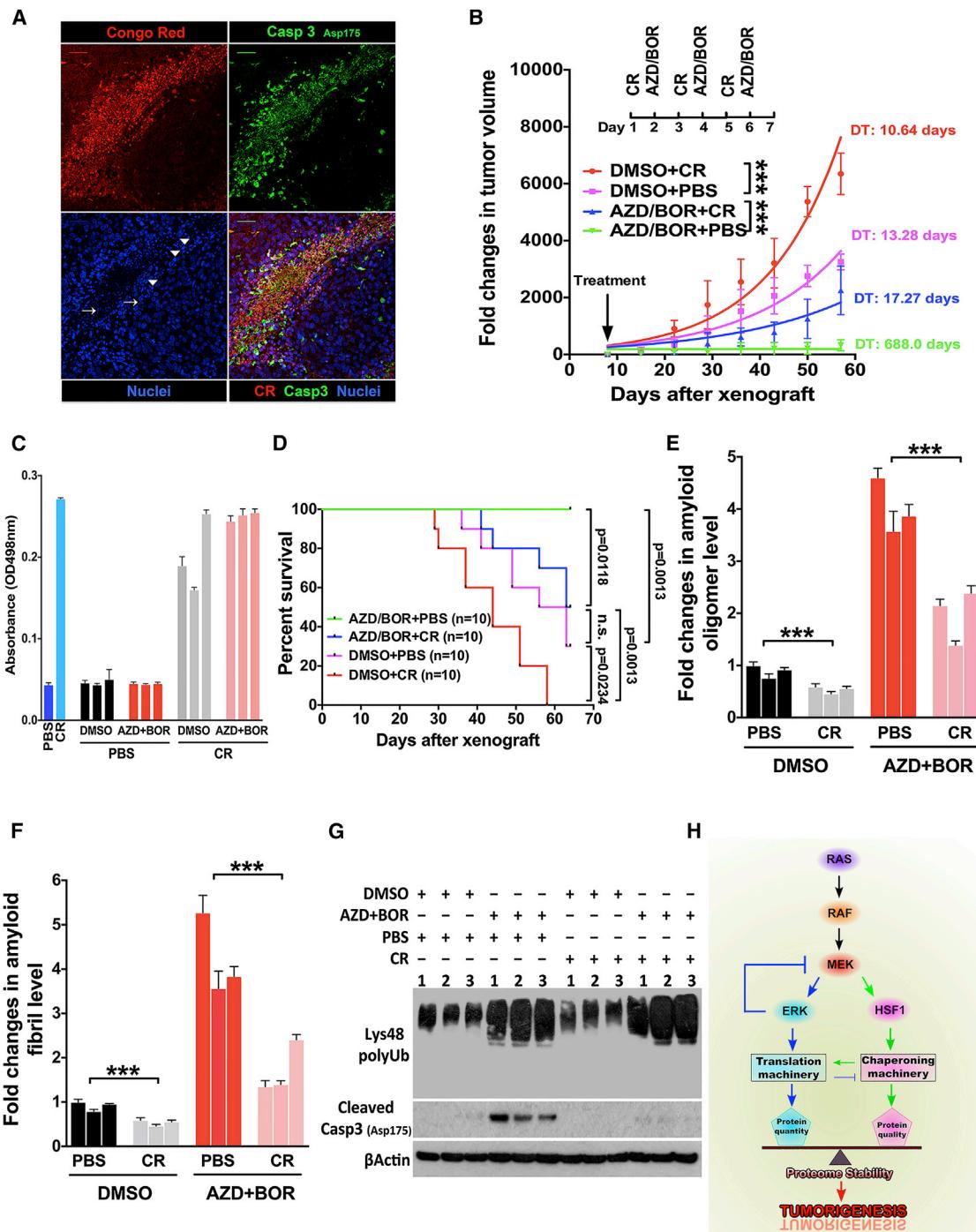
Supplemental Information includes Extended Experimental Procedures, seven figures, and two tables and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2015.01.028>.

## AUTHOR CONTRIBUTIONS

Z.T. and S.D. designed and performed experiments. Y.H. performed statistical analyses. R.D. performed pathological diagnoses. L.S. provided mice and engaged in discussions. C.D. conceived the project and oversaw the studies. S.S. and C.D. wrote the manuscript.

## ACKNOWLEDGMENTS

We thank Luke Whitesell for cell lines, Pete Finger for technical assistance, and members of the C.D. laboratory for discussions. This work was supported



**Figure 7. Amyloidogenesis Suppresses Tumor Growth**

(A) Sections of melanomas receiving combined treatment were stained with cleaved caspase 3 antibodies followed by CR staining. Arrowheads and arrows indicate condensed and fragmented nuclei, respectively. Scale bar represents 50  $\mu$ m.

(B) A2058 cells ( $1 \times 10^6$ ) were s.c. injected into NOD/SCID mice. After 7 days, mice were treated with 1 mg/30 g CR via i.p. injection 1 day prior to combined treatment. Tumor volumes were measured weekly (mean  $\pm$  SD, n = 3). Lysis buffer containing CR served as a positive control.

(D) Kaplan-Meier survival curves were compared (Log-rank test).

(E and F) Both detergent-soluble and detergent-insoluble fractions of tumor lysates were used to quantitate amyloids, three tumors per group (mean  $\pm$  SD, n = 3, Student's t test).

(legend continued on next page)

in part by The Jackson Laboratory Cancer Center Support Grant (3P30CA034196), and grants from NIH (1DP2OD007070) and the Ellison Medical Foundation (AG-NS-0599-09) to C.D.

Received: June 15, 2014

Revised: October 9, 2014

Accepted: December 29, 2014

Published: February 12, 2015

## REFERENCES

Ahn, N.G., Seger, R., Bratlien, R.L., Diltz, C.D., Tonks, N.K., and Krebs, E.G. (1991). Multiple components in an epidermal growth factor-stimulated protein kinase cascade. In vitro activation of a myelin basic protein/microtubule-associated protein 2 kinase. *J. Biol. Chem.* 266, 4220–4227.

Alavez, S., Vantipalli, M.C., Zucker, D.J., Klang, I.M., and Lithgow, G.J. (2011). Amyloid-binding compounds maintain protein homeostasis during ageing and extend lifespan. *Nature* 472, 226–229.

Balch, W.E., Morimoto, R.I., Dillin, A., and Kelly, J.W. (2008). Adapting proteostasis for disease intervention. *Science* 319, 916–919.

Brunet, A., Pagès, G., and Pouysségur, J. (1994a). Constitutively active mutants of MAP kinase kinase (MEK1) induce growth factor-relaxation and oncogenicity when expressed in fibroblasts. *Oncogene* 9, 3379–3387.

Brunet, A., Pagès, G., and Pouysségur, J. (1994b). Growth factor-stimulated MAP kinase induces rapid retrophosphorylation and inhibition of MAP kinase kinase (MEK1). *FEBS Lett.* 346, 299–303.

Catalanotti, F., Reyes, G., Jesenberger, V., Galabova-Kovacs, G., de Matos Simoes, R., Carugo, O., and Baccarini, M. (2009). A Mek1-Mek2 heterodimer determines the strength and duration of the Erk signal. *Nat. Struct. Mol. Biol.* 16, 294–303.

Chiti, F., and Dobson, C.M. (2006). Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem.* 75, 333–366.

Chu, B., Zhong, R., Soncin, F., Stevenson, M.A., and Calderwood, S.K. (1998). Transcriptional activity of heat shock factor 1 at 37 degrees C is repressed through phosphorylation on two distinct serine residues by glycogen synthase kinase 3 and protein kinases Calpha and Czeta. *J. Biol. Chem.* 273, 18640–18646.

Clausson, C.M., Allalou, A., Weibrech, I., Mahmoudi, S., Farnebo, M., Landegren, U., Wählby, C., and Söderberg, O. (2011). Increasing the dynamic range of in situ PLA. *Nat. Methods* 8, 892–893.

Dai, C., Whitesell, L., Rogers, A.B., and Lindquist, S. (2007). Heat shock factor 1 is a powerful multifaceted modifier of carcinogenesis. *Cell* 130, 1005–1018.

Dai, C., Dai, S., and Cao, J. (2012a). Proteotoxic stress of cancer: implication of the heat-shock response in oncogenesis. *J. Cell. Physiol.* 227, 2982–2987.

Dai, C., Santagata, S., Tang, Z., Shi, J., Cao, J., Kwon, H., Bronson, R.T., Whitesell, L., and Lindquist, S. (2012b). Loss of tumor suppressor NF1 activates HSF1 to promote carcinogenesis. *J. Clin. Invest.* 122, 3742–3754.

Dhillon, A.S., Hagan, S., Rath, O., and Kolch, W. (2007). MAP kinase signalling pathways in cancer. *Oncogene* 26, 3279–3290.

Eisenberg, D., and Jucker, M. (2012). The amyloid state of proteins in human diseases. *Cell* 148, 1188–1203.

Favata, M.F., Horiuchi, K.Y., Manos, E.J., Daulerio, A.J., Stradley, D.A., Feeser, W.S., Van Dyk, D.E., Pitts, W.J., Earl, R.A., Hobbs, F., et al. (1998). Identification of a novel inhibitor of mitogen-activated protein kinase kinase. *J. Biol. Chem.* 273, 18623–18632.

Glace, C.G. (2008). Structural classification of toxic amyloid oligomers. *J. Biol. Chem.* 283, 29639–29643.

Guettouche, T., Boellmann, F., Lane, W.S., and Voellmy, R. (2005). Analysis of phosphorylation of human heat shock factor 1 in cells experiencing a stress. *BMC Biochem.* 6, 4.

Hsu, A.L., Murphy, C.T., and Kenyon, C. (2003). Regulation of aging and age-related disease by DAF-16 and heat-shock factor. *Science* 300, 1142–1145.

Isaacs, J.S., Jung, Y.J., Mimnaugh, E.G., Martinez, A., Cuttitta, F., and Neckers, L.M. (2002). Hsp90 regulates a von Hippel Lindau-independent hypoxia-inducible factor-1 alpha-degradative pathway. *J. Biol. Chem.* 277, 29936–29944.

Jin, X., Moskophidis, D., and Mivechi, N.F. (2011). Heat shock transcription factor 1 is a key determinant of HCC development by regulating hepatic steatosis and metabolic syndrome. *Cell Metab.* 14, 91–103.

Kawaguchi, Y., Kovacs, J.J., McLaurin, A., Vance, J.M., Ito, A., and Yao, T.P. (2003). The deacetylase HDAC6 regulates aggresome formation and cell viability in response to misfolded protein stress. *Cell* 115, 727–738.

Kayed, R., Head, E., Thompson, J.L., McIntire, T.M., Milton, S.C., Cotman, C.W., and Glabe, C.G. (2003). Common structure of soluble amyloid oligomers implies common mechanism of pathogenesis. *Science* 300, 486–489.

Kayed, R., Head, E., Sarsota, F., Saing, T., Cotman, C.W., Necula, M., Margol, L., Wu, J., Breydo, L., Thompson, J.L., et al. (2007). Fibril specific, conformation dependent antibodies recognize a generic epitope common to amyloid fibrils and fibrillar oligomers that is absent in prefibrillar oligomers. *Mol. Neurodegener.* 2, 18.

Kim, W., Bennett, E.J., Huttlin, E.L., Guo, A., Li, J., Possemato, A., Sowa, M.E., Rad, R., Rush, J., Comb, M.J., et al. (2011). Systematic and quantitative assessment of the ubiquitin-modified proteome. *Mol. Cell* 44, 325–340.

Laplante, M., and Sabatini, D.M. (2012). mTOR signaling in growth control and disease. *Cell* 149, 274–293.

Lindquist, S. (1986). The heat-shock response. *Annu. Rev. Biochem.* 55, 1151–1191.

Massey, A.J., Williamson, D.S., Browne, H., Murray, J.B., Dokurno, P., Shaw, T., Macias, A.T., Daniels, Z., Geoffroy, S., Dopson, M., et al. (2010). A novel, small molecule inhibitor of Hsc70/Hsp70 potentiates Hsp90 inhibitor induced apoptosis in HCT116 colon carcinoma cells. *Cancer Chemother. Pharmacol.* 66, 535–545.

Meng, L., Gabai, V.L., and Sherman, M.Y. (2010). Heat-shock transcription factor HSF1 has a critical role in human epidermal growth factor receptor-2-induced cellular transformation and tumorigenesis. *Oncogene* 29, 5204–5213.

Min, J.N., Huang, L., Zimonjic, D.B., Moskophidis, D., and Mivechi, N.F. (2007). Selective suppression of lymphomas by functional loss of Hsf1 in a p53-deficient mouse model for spontaneous tumors. *Oncogene* 26, 5086–5097.

Morimoto, R.I. (2008). Proteotoxic stress and inducible chaperone networks in neurodegenerative disease and aging. *Genes Dev.* 22, 1427–1438.

Morris, E.J., Jha, S., Restaino, C.R., Dayananth, P., Zhu, H., Cooper, A., Carr, D., Deng, Y., Jin, W., Black, S., et al. (2013). Discovery of a novel ERK inhibitor with activity in models of acquired resistance to BRAF and MEK inhibitors. *Cancer Discov.* 3, 742–750.

Müller, L., Schaupp, A., Walerych, D., Wegele, H., and Buchner, J. (2004). Hsp90 regulates the activity of wild type p53 under physiological and elevated temperatures. *J. Biol. Chem.* 279, 48846–48854.

Neckers, L., and Workman, P. (2012). Hsp90 molecular chaperone inhibitors: are we there yet? *Clin. Cancer Res.* 18, 64–76.

Nikolaev, S.I., Rimoldi, D., Iseli, C., Valsesia, A., Robyr, D., Gehrig, C., Harshman, K., Guipponi, M., Bukach, O., Zoete, V., et al. (2012). Exome sequencing identifies recurrent somatic MAP2K1 and MAP2K2 mutations in melanoma. *Nat. Genet.* 44, 133–139.

(G) Proteins were detected by immunoblotting, three tumors per group.

(H) Schematic depiction of the interplay among MEK, ERK, and HSF1, and its role in regulating proteome stability. Balanced proteostasis suppresses toxic protein aggregation and amyloidogenesis, thereby facilitating tumorigenesis.

See also [Figure S7](#).

Oerlemans, R., Franke, N.E., Assaraf, Y.G., Cloos, J., van Zantwijk, I., Berkers, C.R., Scheffer, G.L., Debipersad, K., Vojteckova, K., Lemos, C., et al. (2008). Molecular basis of bortezomib resistance: proteasome subunit beta5 (PSMB5) gene mutation and overexpression of PSMB5 protein. *Blood* 112, 2489–2499.

Ohori, M., Kinoshita, T., Okubo, M., Sato, K., Yamazaki, A., Arakawa, H., Nishimura, S., Inamura, N., Nakajima, H., Neya, M., et al. (2005). Identification of a selective ERK inhibitor and structural determination of the inhibitor-ERK2 complex. *Biochem. Biophys. Res. Commun.* 336, 357–363.

Orr, H.T., and Zoghbi, H.Y. (2007). Trinucleotide repeat disorders. *Annu. Rev. Neurosci.* 30, 575–621.

Pickart, C.M., and Eddins, M.J. (2004). Ubiquitin: structures, functions, mechanisms. *Biochim. Biophys. Acta* 1695, 55–72.

Roux, P.P., and Blenis, J. (2004). ERK and p38 MAPK-activated protein kinases: a family of protein kinases with diverse biological functions. *Microbiol. Mol. Biol. Rev.* 68, 320–344.

Sánchez, I., Mahlke, C., and Yuan, J. (2003). Pivotal role of oligomerization in expanded polyglutamine neurodegenerative disorders. *Nature* 421, 373–379.

Sladewski, T.E., Shafer, A.M., and Hoag, C.M. (2006). The effect of ionic strength on the UV-vis spectrum of congo red in aqueous solution. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* 65, 985–987.

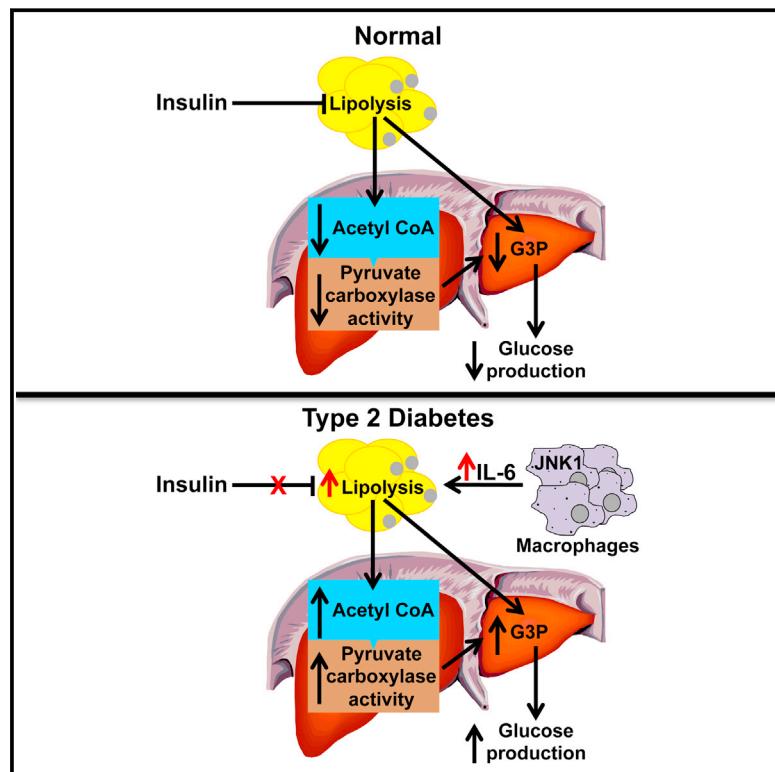
Taipale, M., Jarosz, D.F., and Lindquist, S. (2010). HSP90 at the hub of protein homeostasis: emerging mechanistic insights. *Nat. Rev. Mol. Cell Biol.* 11, 515–528.

Xiao, X., Zuo, X., Davis, A.A., McMillan, D.R., Curry, B.B., Richardson, J.A., and Benjamin, I.J. (1999). HSF1 is required for extra-embryonic development, postnatal growth and protection during inflammatory responses in mice. *EMBO J.* 18, 5943–5952.

Yeh, T.C., Marsh, V., Bernat, B.A., Ballard, J., Colwell, H., Evans, R.J., Parry, J., Smith, D., Brandhuber, B.J., Gross, S., et al. (2007). Biological characterization of ARRY-142886 (AZD6244), a potent, highly selective mitogen-activated protein kinase kinase 1/2 inhibitor. *Clin. Cancer Res.* 13, 1576–1583.

# Hepatic Acetyl CoA Links Adipose Tissue Inflammation to Hepatic Insulin Resistance and Type 2 Diabetes

## Graphical Abstract



## Authors

Rachel J. Perry,  
João-Paulo G. Camporez, ...,  
Kitt Falk Petersen, Gerald I. Shulman

## Correspondence

gerald.shulman@yale.edu

## In Brief

Metabolic abnormalities associated with a high-fat diet are found to be driven by increased hepatic acetyl CoA levels, which are shown to be a consequence of white adipose tissue inflammation and inappropriately increased lipolysis.

## Highlights

- Insulin inhibits gluconeogenesis by suppressing lipolysis and hepatic acetyl CoA
- Hyperglycemia associated with HFD is due to increased WAT-derived hepatic acetyl CoA
- ATGL KOs are protected from HFD-induced insulin resistance due to decreased lipolysis
- mφJNK KOs are protected from HFD-induced insulin resistance due to decreased lipolysis

# Hepatic Acetyl CoA Links Adipose Tissue Inflammation to Hepatic Insulin Resistance and Type 2 Diabetes

Rachel J. Perry,<sup>1,2,3</sup> João-Paulo G. Camporez,<sup>2</sup> Romy Kursawe,<sup>2</sup> Paul M. Titchenell,<sup>4</sup> Dongyan Zhang,<sup>1</sup> Curtis J. Perry,<sup>5</sup> Michael J. Jurczak,<sup>2</sup> Abulizi Abudukadier,<sup>2</sup> Myoung Sook Han,<sup>6</sup> Xian-Man Zhang,<sup>1</sup> Hai-Bin Ruan,<sup>7</sup> Xiaoyong Yang,<sup>3,7</sup> Sonia Caprio,<sup>8</sup> Susan M. Kaech,<sup>5</sup> Hei Sook Sul,<sup>9</sup> Morris J. Birnbaum,<sup>4</sup> Roger J. Davis,<sup>1,6</sup> Gary W. Cline,<sup>2</sup> Kitt Falk Petersen,<sup>2</sup> and Gerald I. Shulman<sup>1,2,3,\*</sup>

<sup>1</sup>Howard Hughes Medical Institute, Yale University, New Haven, CT 06519, USA

<sup>2</sup>Department of Internal Medicine, Yale University, New Haven, CT 06520, USA

<sup>3</sup>Department of Cellular and Molecular Physiology, Yale University, New Haven, CT 06520, USA

<sup>4</sup>The Institute for Diabetes, Obesity and Metabolism, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

<sup>5</sup>Department of Immunobiology, Yale University, New Haven, CT 06520, USA

<sup>6</sup>Program in Molecular Medicine, University of Massachusetts Medical School, Worcester, MA 01605, USA

<sup>7</sup>Department of Comparative Medicine, Yale University, New Haven, CT 06520, USA

<sup>8</sup>Department of Pediatrics, Yale University, New Haven, CT 06520, USA

<sup>9</sup>Department of Nutritional Sciences and Toxicology, University of California, Berkeley, Berkeley, CA 94720, USA

\*Correspondence: [gerald.shulman@yale.edu](mailto:gerald.shulman@yale.edu)

<http://dx.doi.org/10.1016/j.cell.2015.01.012>

## SUMMARY

Impaired insulin-mediated suppression of hepatic glucose production (HGP) plays a major role in the pathogenesis of type 2 diabetes (T2D), yet the molecular mechanism by which this occurs remains unknown. Using a novel *in vivo* metabolomics approach, we show that the major mechanism by which insulin suppresses HGP is through reductions in hepatic acetyl CoA by suppression of lipolysis in white adipose tissue (WAT) leading to reductions in pyruvate carboxylase flux. This mechanism was confirmed in mice and rats with genetic ablation of insulin signaling and mice lacking adipose triglyceride lipase. Insulin's ability to suppress hepatic acetyl CoA, PC activity, and lipolysis was lost in high-fat-fed rats, a phenomenon reversible by IL-6 neutralization and inducible by IL-6 infusion. Taken together, these data identify WAT-derived hepatic acetyl CoA as the main regulator of HGP by insulin and link it to inflammation-induced hepatic insulin resistance associated with obesity and T2D.

## INTRODUCTION

The Centers for Disease Control predict that by the year 2050, one in three Americans will suffer from type 2 diabetes (T2D) (Boyle et al., 2010); therefore, efforts to understand and intervene in the pathogenesis of the disease are urgently needed. Fasting hyperglycemia, due to increased rates of hepatic gluconeogenesis, and the inability of insulin to suppress this process are the hallmark of T2D, but the molecular mechanism for this alteration is unknown (Kahn et al., 2014; Olefsky and Glass, 2010; Robbins

et al., 2014; Samuel and Shulman, 2012). In this regard, mice lacking Akt1, Akt2, and Foxo1 (TKO) (Lu et al., 2012), the canonical insulin signaling proteins thought to regulate hepatic glucose production (HGP), surprisingly do not show any defect in insulin-mediated suppression of HGP.

We hypothesized that insulin's ability to suppress lipolysis in white adipose tissue (WAT) may be critical for the suppression of HGP by reducing fatty acid flux to the liver resulting in decreased hepatic acetyl CoA concentrations and decreased pyruvate carboxylase (PC) activity, the first step in the conversion of pyruvate to glucose. While hepatic acetyl CoA has been previously shown to modulate PC activity *in vitro* (Barritt et al., 1966; Cazzulo and Stoppani, 1968; Cooper and Benedict, 1966; Keech and Utter, 1963; Krebs et al., 1965), methodological limitations stemming from acetyl CoA's low concentrations and rapid degradation *ex vivo* have prevented its measurement *in vivo*.

We further hypothesized that increased hepatic acetyl CoA, due to increased rates of WAT lipolysis, is responsible for increased fasting HGP and reduced insulin-mediated suppression of HGP in high-fat-fed rodent models of diet-induced obesity associated with hepatic insulin resistance and hyperglycemia. In this regard, inflammatory cytokines are well-established mediators of lipolysis (Fonseca-Alaniz et al., 2007; Savage et al., 2005; Suganami and Ogawa, 2010), and protection from inflammation-associated insulin resistance could explain the preserved insulin sensitivity of macrophage-specific c-JUN N-terminal kinase ( $m_{\phi}$ JNK) knockout mice fed a high-fat diet (Han et al., 2013).

To address these questions, we developed a novel liquid chromatography/tandem mass spectrometry (LC-MS/MS) technique to measure hepatic acetyl CoA content from *in situ* freeze-clamped liver within 10 s of intravenous euthanasia in awake, unrestrained rats. These measurements were combined with measurements of whole-body rates of lipolysis, assessed by

rates of [ $1,1,2,3,3\text{-}^2\text{H}_5$ ] glycerol and [ $\text{U-}^{13}\text{C}$ ] palmitate turnover, combined with rates of glucose turnover assessed by [ $3\text{-}^3\text{H}$ ] glucose before and during a hyperinsulinemic-euglycemic clamp. Finally, we also performed additional studies to quantify *in vivo* pyruvate carboxylase flux and relative contributions of oxaloacetate and glycerol to hepatic gluconeogenesis using a combined NMR-LC-MS/MS positional isotopomer approach with [ $3\text{-}^{13}\text{C}$ ] lactate as a tracer. Using this comprehensive *in vivo* metabolomics approach, we show that hepatic acetyl CoA is the key regulator of hepatic insulin action in normal fasted, free-ranging rodents *in vivo* and that fasting hyperglycemia and insulin resistance in high-fat-fed rodents result from an inappropriate increase in hepatic acetyl CoA concentrations due to macrophage-induced WAT lipolysis.

## RESULTS

### Insulin Suppression of Glucose Production Is Temporally Associated with Suppression of Lipolysis and Hepatic Acetyl CoA

In order to examine the respective roles of hepatic acetyl CoA and glycerol turnover in the regulation of HGP by insulin, we infused awake rats with insulin to mimic physiologic postprandial insulin concentrations during a hyperinsulinemic-euglycemic clamp and found that plasma non-esterified fatty acid (NEFA) and glycerol concentrations were suppressed by 90% within 5 min, and HGP was suppressed by ~70% within 10 min after initiation of the insulin infusion as reflected by a rapid increase in the plasma glucose specific activity (Figures 1A–1C; Figures S1A–S1D). This reduction in HGP was associated with a 50% reduction in hepatic acetyl CoA concentrations within 10 min of the start of the insulin infusion as well as a 60% suppression of whole-body glycerol and palmitate turnover at steady state with no changes in plasma lactate, glucagon, or liver glycogen concentrations (Figures 1D–1F and S1E–S1G). Therefore, these data demonstrate that the major effect of insulin to suppress HGP under these conditions was through suppression of hepatic gluconeogenesis.

### Clamping Hepatic Acetyl CoA and Whole-Body Glycerol Turnover Abrogates Insulin Suppression of HGP

Based on the previous data, we next examined whether altering hepatic acetyl CoA concentrations, and/or altering rates of whole-body glycerol turnover, would change rates of HGP during a hyperinsulinemic-euglycemic clamp. In order to modify hepatic acetyl CoA concentrations, we infused sodium acetate at increasing rates during a hyperinsulinemic-euglycemic clamp, and observed a strong correlation between hepatic acetyl CoA concentrations and both pyruvate carboxylase (PC) activity ( $r^2 = 0.85$ ,  $p < 0.0001$ ) and HGP ( $r^2 = 0.68$ ,  $p < 0.0001$ ) during the clamp, without any changes in liver glycogen content (Figures 1G–1I and S1H–S1M).

Likewise, in order to assess the potential contribution of insulin-mediated suppression of glycerol turnover to insulin-mediated suppression of HGP by a substrate-push mechanism, we infused glycerol at increasing rates in awake rats during a hyperinsulinemic-euglycemic clamp and observed a strong correlation ( $r^2 = 0.55$ ,  $p < 0.0001$ ) between rates of glycerol turnover

and rates of HGP, without any change in liver glycogen or acetyl CoA (Figures 1J–1L and S1N–S1R). Taken together, these data demonstrate that modulation of either hepatic acetyl CoA concentrations or glycerol turnover can independently modulate rates of HGP during a hyperinsulinemic-euglycemic clamp.

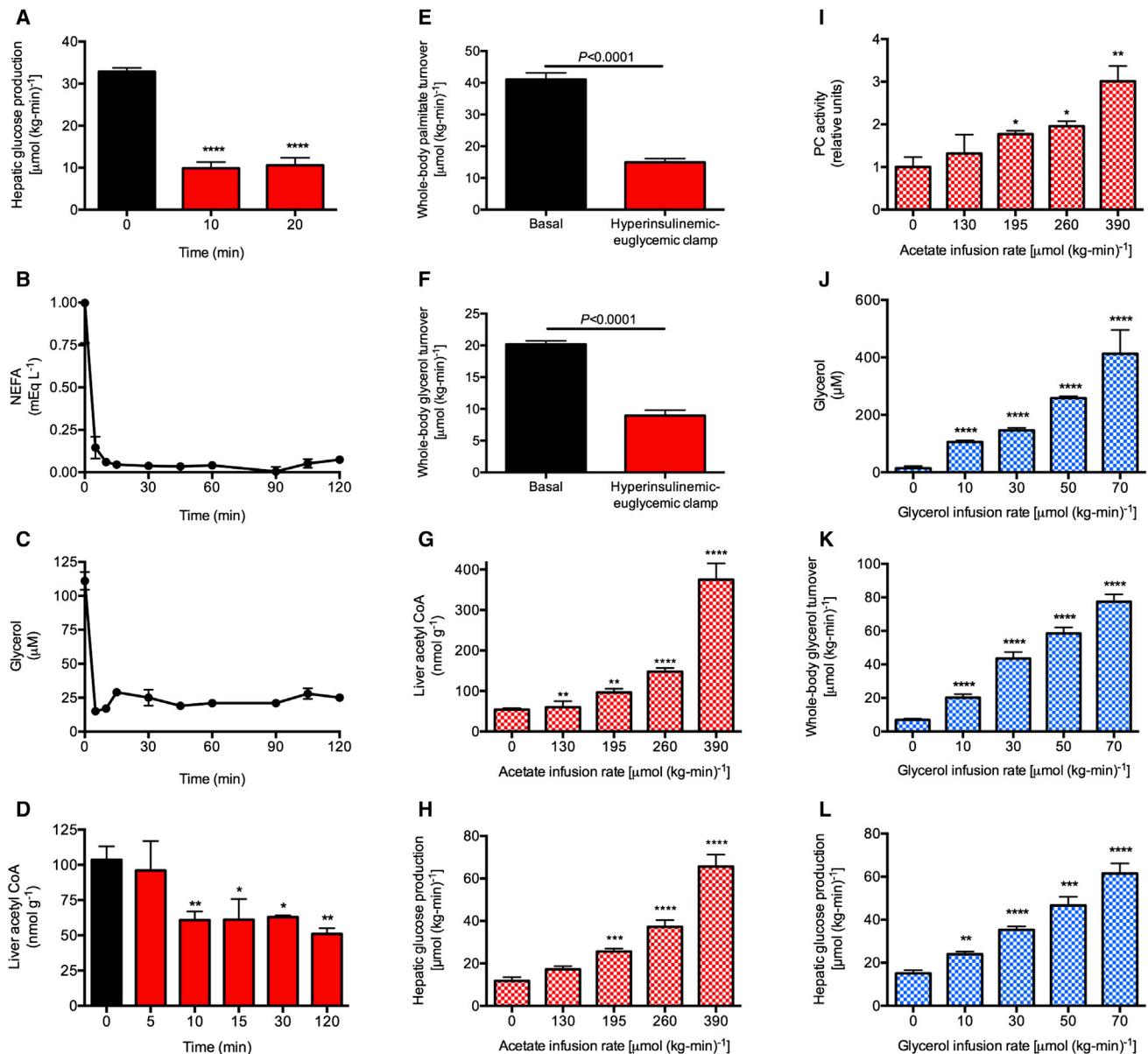
In order to determine whether insulin-mediated suppression of hepatic acetyl CoA concentrations and glycerol turnover were responsible for insulin-mediated suppression of HGP, we co-infused both acetate and glycerol to match hepatic acetyl CoA concentrations and rates of glycerol turnover during a hyperinsulinemic-euglycemic clamp to those observed under basal conditions (Figures 2A–2D; Figures S2A and S2B). Matching hepatic acetyl CoA concentrations and whole-body glycerol turnover rates to baseline levels completely abrogated insulin-mediated suppression of PC flux and HGP, without any changes to plasma glucagon or liver glycogen concentrations, or expression of gluconeogenic proteins in the liver (Figures 2E, 2F, and S2C–S2E). Multiple regression analysis confirmed the linear relationships between HGP and both liver acetyl CoA and whole-body glycerol turnover ( $p = 0.002$  and  $p = 0.01$ , respectively). We next assessed the relative contributions of oxaloacetate and glycerol conversion to glucose before and during a clamp and found that the suppression of hepatic gluconeogenesis could entirely be ascribed to suppression of lipolysis, with 85% of the suppression of HGP attributed to reduced conversion of oxaloacetate to glucose and the remainder attributed to reduced conversion of glycerol to glucose (Figure 2E).

### Liver Akt1/Ak2/Foxo1 Knockout Mice Suppress Lipolysis, Hepatic Acetyl CoA, and HGP Normally, but Acetate and Glycerol Infusion Abrogates Insulin Suppression of HGP

TKO mice exhibited no alterations in basal or insulin-mediated reductions in rates of HGP before and during a hyperinsulinemic-euglycemic clamp despite clear knockdown of Akt1/Akt2 and Foxo1 protein concentrations in liver (Lu et al., 2012) (Figure 3A; Figures S3A–S3C). Consistent with our data in awake rats, hepatic acetyl CoA concentrations, PC activity, and rates of whole-body lipolysis were similarly suppressed in WT and TKO mice during the hyperinsulinemic-euglycemic clamp, without any change in hepatic glycogen content (Figures 3B–3D and S3D–S3G). Furthermore, co-infusion of acetate and glycerol to maintain basal hepatic acetyl CoA concentrations and basal rates of glycerol turnover, fully abrogated insulin's ability to suppress HGP during the hyperinsulinemic-euglycemic clamp (Figures 3B–3D).

### Inhibition of Lipolysis Decreases Hepatic Acetyl CoA and HGP in Insulin Receptor Antisense Oligonucleotide-Treated Rats Lacking Insulin Receptor in Liver and Fat

In order to more directly examine the putative requirement for suppression of lipolysis to permit suppression of HGP, we treated rats with a 2'-O-methoxyethyl chimeric antisense oligonucleotide to knock down the insulin receptor (IR ASO) in liver and adipose tissue (Figure 3E). IR ASO-treated rats displayed severe insulin resistance manifested by failure to suppress lipolysis, hepatic acetyl CoA, and HGP during a hyperinsulinemic-euglycemic clamp; however, all of these effects were abrogated



**Figure 1. Rapid Suppression of Hepatic Glucose Production Rates Is Temporally Associated with a Parallel Rapid Suppression of Lipolysis**  
 (A) Hepatic glucose production measured by the Steele equation (Steele, 1959). Comparisons by ANOVA with Bonferroni's multiple comparisons test. \*\*\* $p < 0.0001$  versus time 0.

(B and C) Plasma non-esterified fatty acid (NEFA) and glycerol concentrations during the hyperinsulinemic-euglycemic clamp. In (A)-(C),  $n = 6$  per time point.

(D) Hepatic acetyl CoA concentrations.  $n = 4$  per time point. \* $p < 0.05$ , \*\* $p < 0.01$  versus time 0 by the two-tailed unpaired Student's t test.

(E and F) Whole-body palmitate and glycerol turnover rates.

(G–I) Liver acetyl CoA concentrations, hepatic glucose production, and pyruvate carboxylase activity in rats undergoing a hyperinsulinemic-euglycemic clamp with co-infusion of acetate. In (G)–(L),  $n = 4$  per group.

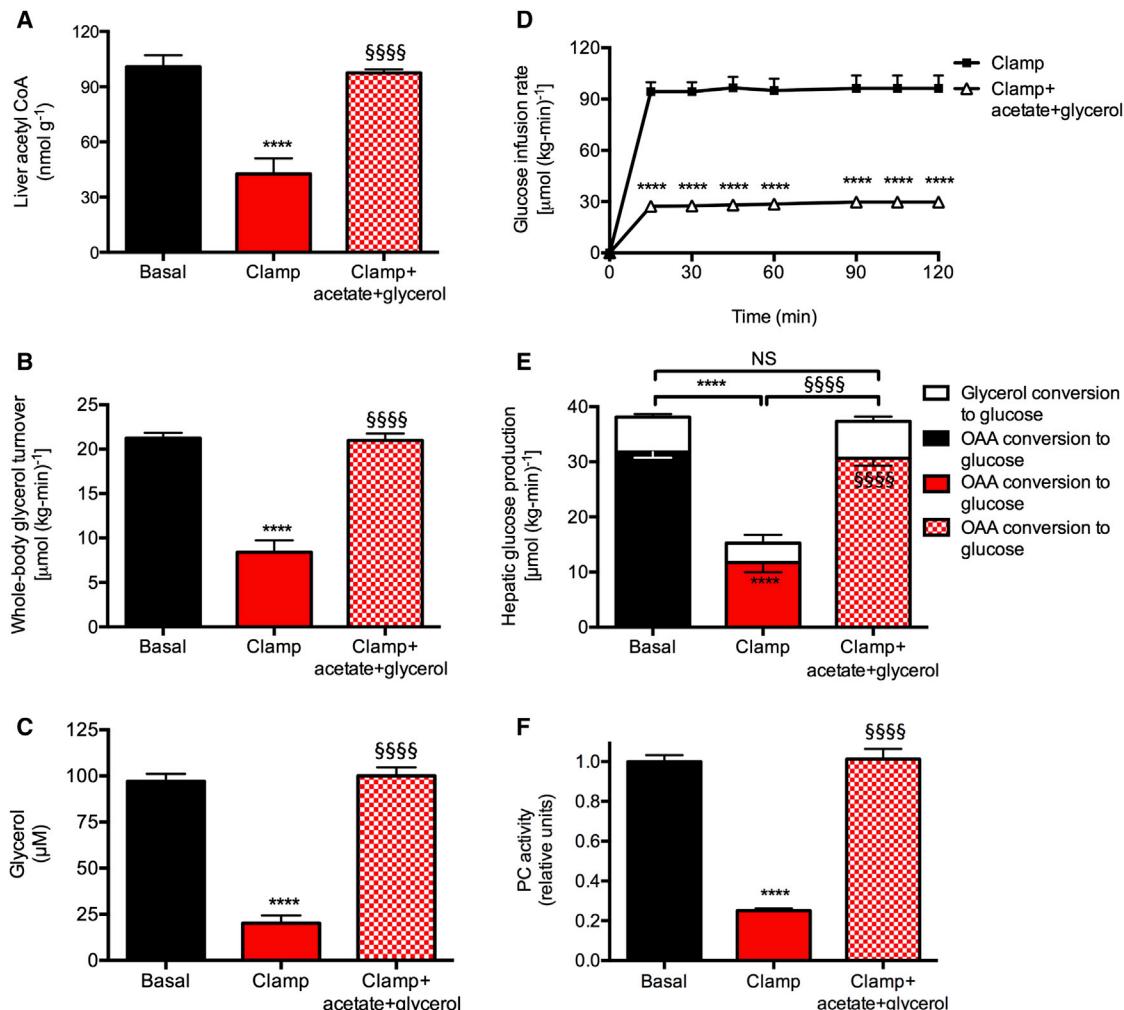
(J–L) Plasma glycerol concentrations, whole-body glycerol turnover, and hepatic glucose production in rats undergoing a hyperinsulinemic-euglycemic clamp with co-infusion of glycerol.

In (G)–(L),  $n = 4$  per group. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , \*\*\*\* $p < 0.0001$  versus 0  $\mu\text{mol}/(\text{kg}\cdot\text{min})$ . In all panels, data are mean  $\pm$  SEM. See also Figure S1.

by administration of atglistatin, a specific inhibitor of adipose triglyceride lipase (Figures 3F–3H and S3H–S3L), attributing the suppression of HGP to the inhibition of lipolysis and subsequent reductions in hepatic acetyl CoA and PC activity in both control and IR knockdown rats.

#### High-Fat-Fed Rats Manifest Increased Rates of HGP and Hepatic Insulin Resistance Associated with Increased Lipolysis and Increased Hepatic Acetyl CoA

In order to understand the potential role of hepatic acetyl CoA in the pathogenesis of fasting and postprandial



**Figure 2. Normalizing Hepatic Acetyl CoA Concentrations and Whole-Body Glycerol Turnover to Baseline Levels in the Hyperinsulinemic-Euglycemic Clamp Abrogates Insulin's Ability to Suppress Hepatic Glucose Production**

(A) Liver acetyl CoA content.

(B) Whole-body glycerol turnover.

(C) Plasma glycerol concentrations.

(D) Glucose infusion rate.

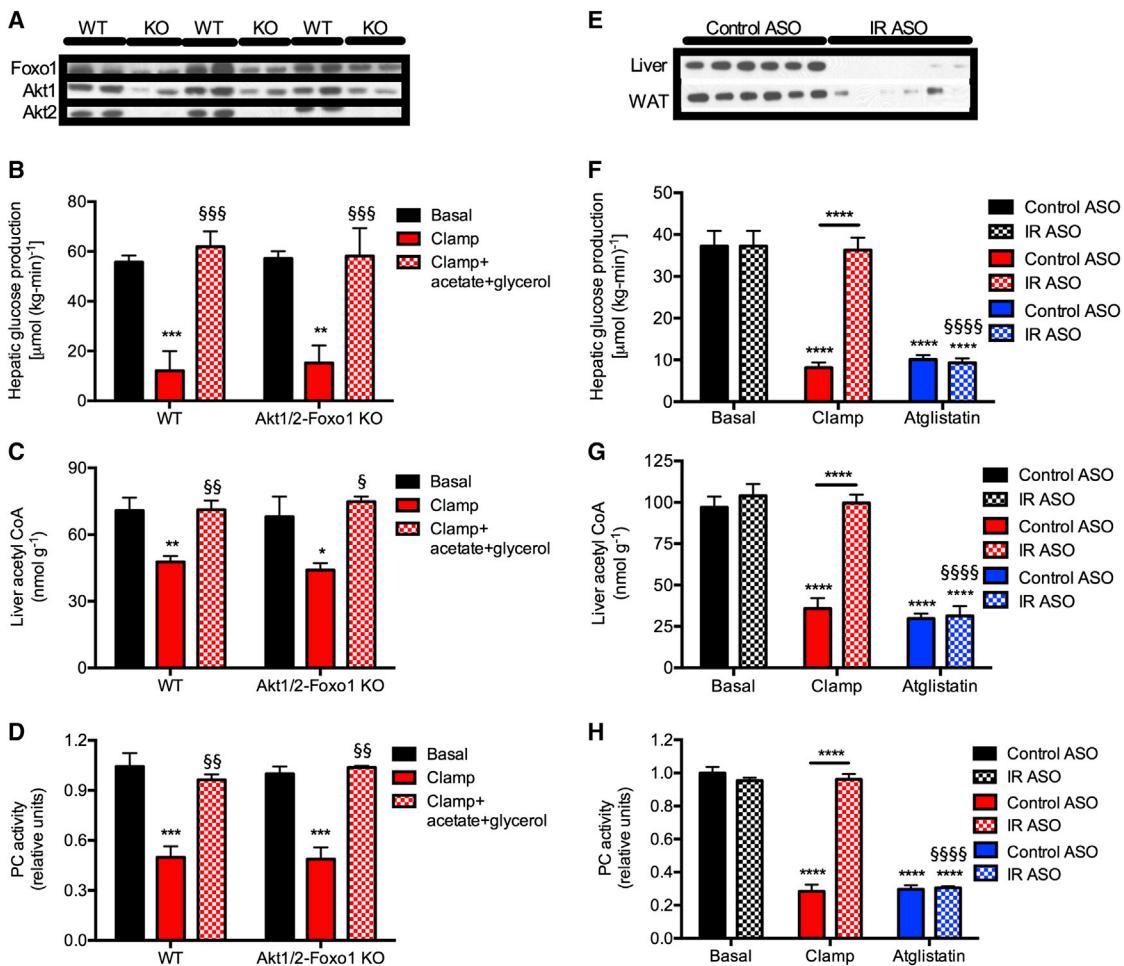
(E) Hepatic glucose production from oxaloacetate (lower bars) and glycerol (upper bars).

(F) PC activity.

In all panels, \*\*\*p < 0.0001 versus basal, §§§§p < 0.0001 versus hyperinsulinemic-euglycemic clamp. n = 6 per group. Data are mean ± SEM. Comparisons by ANOVA with Bonferroni's multiple comparisons test. See also Figure S2.

hyperglycemia in a chronic high-fat-fed rodent model of insulin resistance associated with inflammation, we fed male Sprague-Dawley rats a high-fat diet for 4 weeks and observed, as expected, increases in fasting plasma glucose, insulin, and non-esterified fatty acid (NEFA) concentrations without any increases in plasma glucagon concentrations (Figures S4A–S4D). Fasting hyperglycemia in the HFD group was associated with a 25% increase in rates of basal HGP, 60% of which was derived from increases in hepatic pyruvate carboxylase flux, and 40% from increased conversion of glycerol to glucose. In chow-fed rats, insulin suppressed

HGP in a hyperinsulinemic-euglycemic clamp, 85% through reductions in flux through hepatic pyruvate carboxylase, and 15% through reductions in the contribution of glycerol to glucose (Figure 4A; Figures S4E–S4G). However, in the high-fat-fed rats, insulin failed to suppress HGP during a hyperinsulinemic-euglycemic clamp associated with increased basal and clamped acetyl CoA and PC activity, and with higher rates of lipolysis both before and during the hyperinsulinemic-euglycemic clamp in the HFD group, without any differences in gluconeogenic protein expression (Figures 4B–4E, S4H, and S4I).



**Figure 3. Suppression of Lipolysis Results in Decreased Hepatic Acetyl CoA Concentrations and Suppression of Hepatic Glucose Production Independent of Hepatic Insulin Signaling**

(A) Protein expression in Akt1 and 2 and Foxo1 knockout mice.

(B) Hepatic glucose production. In all panels, black bars, basal; white bars, clamp; and gray bars, clamp+acetate+glycerol.

(C) Hepatic acetyl CoA concentrations.

(D) Hepatic pyruvate carboxylase activity. In (B)–(D), \*p < 0.05, \*\*p < 0.01, \*\*\*p < 0.001 versus basal, §p < 0.05, §§p < 0.01, §§§p < 0.001 versus hyperinsulinemic-euglycemic clamp by ANOVA. Data are mean ± SEM of n = 5 per group (basal), 8–9 per group (hyperinsulinemic-euglycemic clamp), and 3–4 per group (hyperinsulinemic-euglycemic clamp + acetate + glycerol).

(E) Protein expression in insulin receptor knockdown rats.

(F) Hepatic glucose production. Atglistatin-treated rats were given 600  $\mu\text{mol}/\text{kg}$ .

(G) Liver acetyl CoA.

(H) Hepatic PC activity.

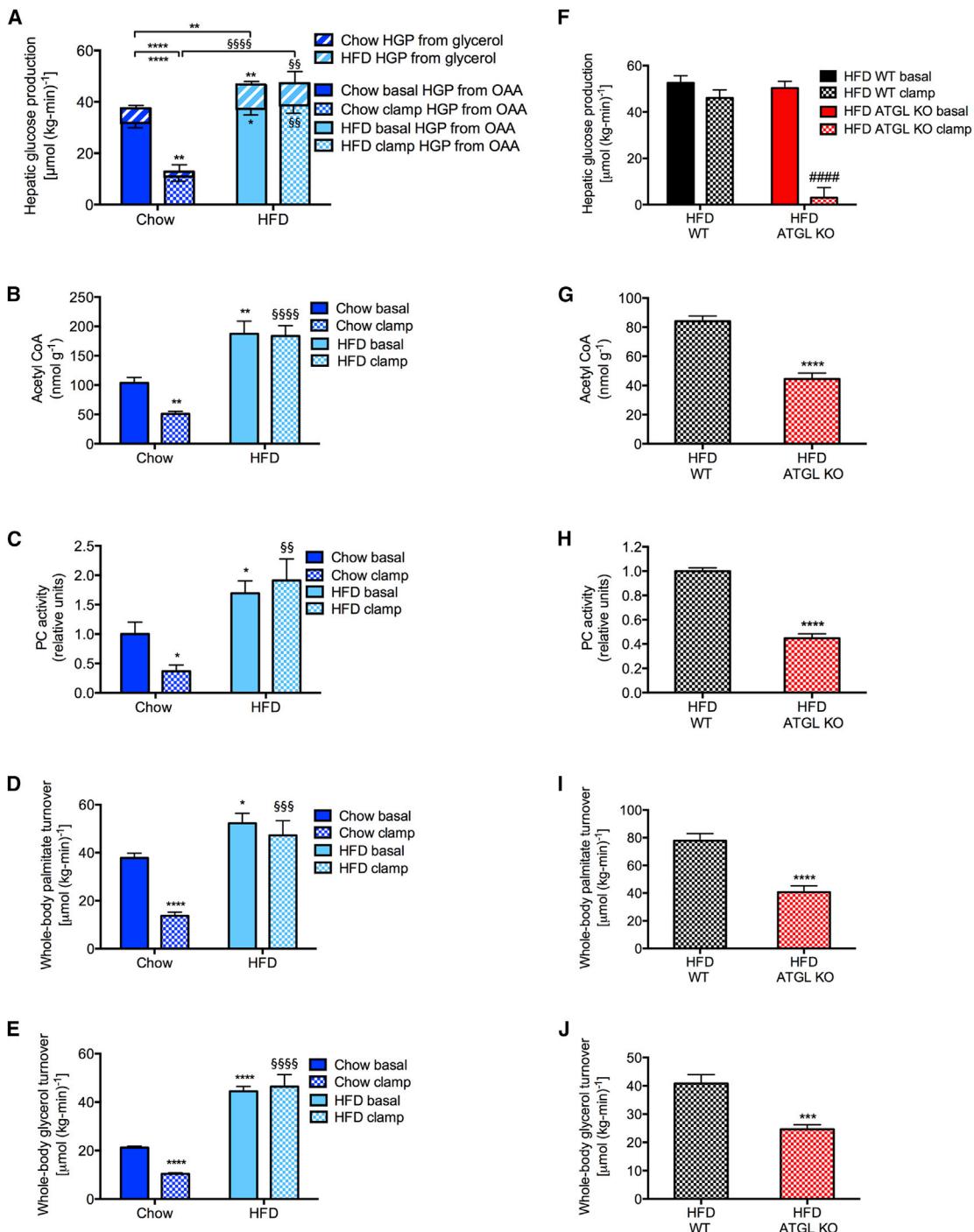
In (F)–(H), \*\*\*\*p < 0.0001 versus basal, §§§§p < 0.0001 versus hyperinsulinemic-euglycemic clamp by ANOVA. Data are mean ± SEM of n = 6 per group. See also Figure S3.

### Increased Lipolysis Plays a Key Role in Excess Hepatic Gluconeogenesis in High-Fat-Fed Rodents

In order to confirm whether excess lipolysis causes excess HGP in diet-induced obesity, we studied high-fat-fed adipose-specific adipose triglyceride lipase (ATGL) knockout mice. These mice exhibited a 5-fold increase in insulin's ability to suppress HGP in a hyperinsulinemic-euglycemic clamp, associated with 40%–60% reductions in hepatic acetyl CoA, PC activity, and whole-body palmitate and glycerol turnover during the clamp (Figures 4F–4J; Figures S5J–S5M).

### High-Fat-Fed Rats Manifest Insulin Resistance in White Adipose Tissue

To determine the primary source(s) of the excess lipolysis in high-fat-fed rodents, we measured the ratio of tissue to plasma  $^{13}\text{C}_{16}$  palmitate in rats infused with  $^{13}\text{C}_{16}$  palmitate. The ratio of tissue to plasma  $^{13}\text{C}_{16}$  palmitate was ~100% in liver, skeletal muscle, and kidney, but only ~65% in WAT, confirming WAT as the source of the majority of fatty acids derived from lipolysis in both the basal and clamped state (Figures S4N and S4O). In addition, in order to determine the proportion of acetyl CoA derived



**Figure 4. Four Weeks of High-Fat Feeding Increases Hepatic Glucose Production as a Result of Increased Lipolysis**

(A) Hepatic glucose production from oxaloacetate and glycerol.

(B and C) Liver acetyl CoA and PC activity.

(D and E) Whole-body palmitate and glycerol turnover.

In (A)–(E), data are mean  $\pm$  SEM of  $n = 6$  per group. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.0001$  versus chow-fed, basal rats by the two-tailed unpaired Student's t test.

(F) Hepatic glucose production in basal and clamped high-fat-fed adipose-specific ATGL knockout mice.

(G and H) Liver acetyl CoA and PC activity.

(I and J) Whole-body palmitate and glycerol turnover.

In (F)–(J), data are mean  $\pm$  SEM of  $n = 8$  per group. \*\*\* $p < 0.001$ , \*\*\*\* $p < 0.0001$  versus basal WT, ##### $p < 0.0001$  versus basal ATGL knockouts by the two-tailed unpaired Student's t test. See also Figures S4 and S5.

from palmitate, we measured the ratio of hepatic  $^{13}\text{C}_2$  acetyl CoA to  $^{13}\text{C}_{16}$  palmitate in rats infused with  $^{13}\text{C}_{16}$  palmitate. In both chow and high-fat-fed rats, under both basal and clamped conditions, the enrichment of hepatic acetyl CoA was approximately 75% of palmitate, showing that a large proportion of hepatic acetyl CoA is derived from palmitate in vivo (Figure S4P).

WAT insulin resistance in high-fat-fed rats was reflected in the failure of insulin to stimulate Akt phosphorylation, glucose uptake, or lipogenesis in WAT (Figures S4Q–S4S). The increased lipolysis was associated with a larger mean diameter of large adipocytes, a reduced number of large adipocytes (Kursawe et al., 2010), and a doubling in WAT comparative gene identification-58 (CGI-58) protein expression (Yamaguchi et al., 2004, 2007), without any other changes in lipolytic or insulin signaling proteins in WAT, liver glycogen or ceramide content, plasma lactate, FGF-21, adiponectin, or lipolytic hormones (Figures S4T–S4EE). Liver concentrations of other putative mediators of gluconeogenesis, including the ratios of ATP to ADP, ATP to AMP, NAD $^+$  to NADH, and four TCA cycle intermediates were also unchanged in high-fat-fed or clamped rats (Figures S5A–S5G).

### Inhibition of Adipose Triglyceride Lipase Normalizes Hepatic Acetyl CoA Content, PC Activity, HGP, and Plasma Glucose Concentrations in High-Fat-Fed Rats

To determine whether increased rates of WAT lipolysis leading to increased hepatic acetyl CoA concentrations and PC activity in HFD rats were responsible for the increased rates of HGP in the basal state and the failure of insulin to suppress HGP during the hyperinsulinemic-euglycemic clamp, we treated rats with atglistatin. Atglistatin treatment normalized fasting plasma glucose and insulin concentrations and HGP by lowering WAT lipolysis, liver acetyl CoA concentrations, and PC activity without affecting plasma glucagon concentrations (Figures S5H–S5Q). These data demonstrate a critical but rapidly reversible role for WAT lipolysis driving increases in hepatic acetyl CoA and PC activity resulting in hyperglycemia in high-fat-fed, insulin-resistant rats.

To further examine the role of increases in hepatic acetyl CoA in causing excess gluconeogenesis, we treated chow and high-fat-fed rats with etomoxir, a potent inhibitor of carnitine palmitoyltransferase-1 (CPT-1). Consistent with the previous studies demonstrating a strong relationship between hepatic acetyl CoA concentration and rates of HGP, etomoxir treatment resulted in reductions in hepatic acetyl CoA content, which was associated with lower fasting plasma glucose and insulin concentrations, reduced rates of HGP, and reduced PC activity without any change in liver glycogen content (Figures S5R–S5W). These data further support the hypothesis that hepatic acetyl CoA is a key regulator of hepatic gluconeogenesis and HGP in normal, fasting rodents.

### Adipose Tissue Inflammation Increases Hepatic Acetyl CoA Content and PC Activity and Causes Hepatic Insulin Resistance in HFD Rats

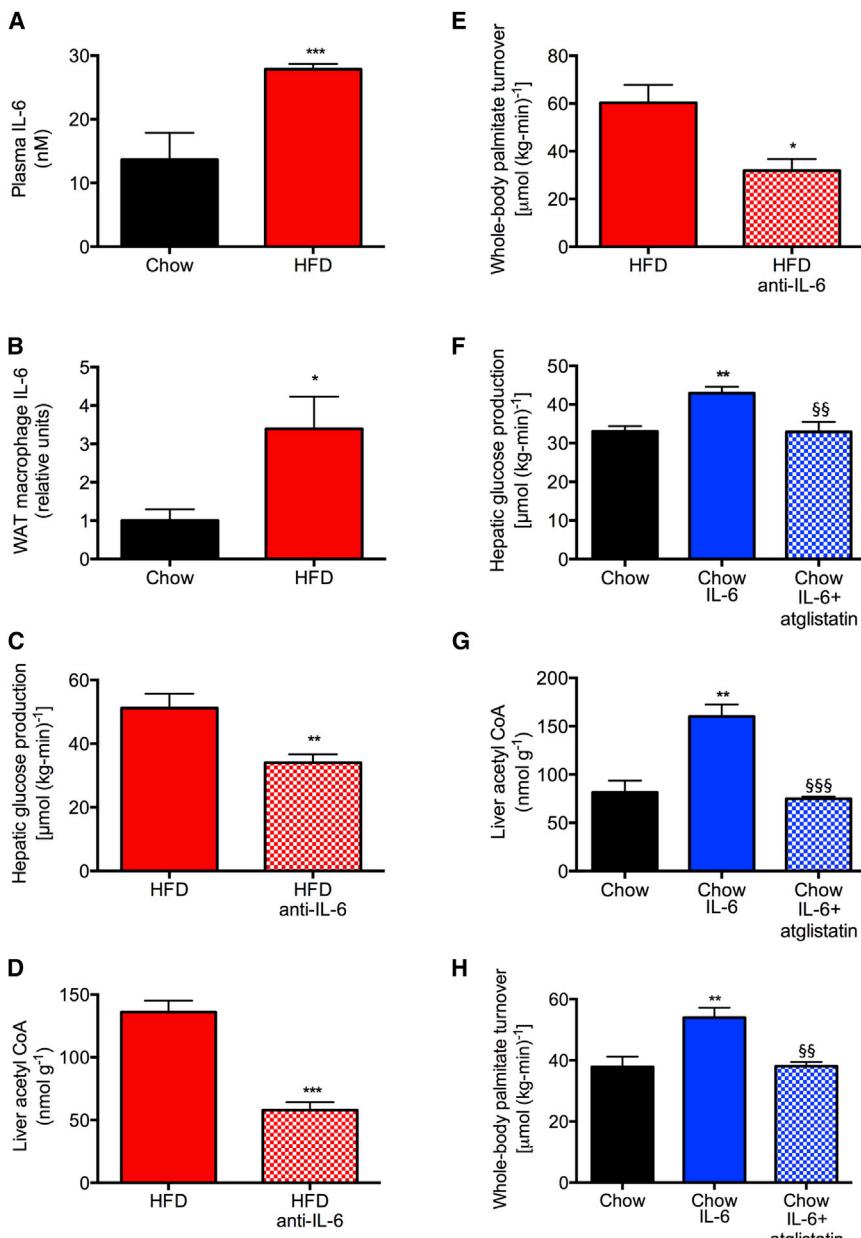
Next, we sought to determine the reason for the increased rates of lipolysis in HFD rats. In the absence of any changes in lipolytic hormone (growth hormone, epinephrine, norepinephrine, corticosterone) concentrations, we hypothesized that adipose tissue

inflammation may play an important role in this process (Mantteiga et al., 2013; Olefsky and Glass, 2010; Robbins et al., 2014). Consistent with this hypothesis, we observed a doubling of plasma and WAT macrophage interleukin-6 (IL-6) concentrations, without changes in any other inflammatory mediators (Figures 5A, 5B, and S6A). We confirmed white adipose tissue as a target of IL-6 by incubating WAT with the concentrations of IL-6 measured in plasma in vivo (15 and 25 nM) and observed a 50% increase in fatty acid release into the media over 24 hr (Figure S6B). Increases in granulocyte macrophage colony stimulating factor (GM-CSF) in both plasma and WAT may be responsible for attracting increased macrophages to WAT in high-fat-fed rats and/or inducing their proliferation, and IL-6 has been shown to stimulate lipolysis through activation of the JAK-Stat signaling pathway (Cernkovich et al., 2008; Guschin et al., 1995; Lütticken et al., 1994; Stahl et al., 1994). A modest 20% increase in WAT tumor necrosis factor- $\alpha$  (TNF- $\alpha$ ) was also observed, consistent with increased macrophage activation in HFD rats (Figures S6C–S6D). Time-course studies with varying periods of HFD treatment confirmed a temporal relationship between fasting plasma glucose, insulin, and NEFA concentrations and both plasma and WAT macrophage IL-6 and GM-CSF concentrations (Figures S6E–S6K), with each increasing between days 6 and 8. In addition, by 6 days of high-fat feeding, adipocytes had begun to exhibit hypertrophy (Figure S6L).

In order to directly assess the role of excess IL-6 in promoting increased lipolysis, we treated HFD rats with a monoclonal neutralizing antibody to IL-6 (Figure S6M). Neutralizing IL-6 in HFD rats reduced plasma glucose, NEFA, glycerol, and insulin concentrations to normal levels without any change in plasma glucagon concentrations (Figures S6N–S6S). By lowering rates of lipolysis and reducing hepatic acetyl CoA concentrations, anti-IL-6 treatment resulted in a 50% reduction in hepatic PC activity and normalized rates of HGP in the HFD rat, without affecting macrophage TNF- $\alpha$  concentrations (Figures 5C–5F and S6T–S6V).

### IL-6 Infusion Induces Hepatic Insulin Resistance through Increased Adipose Tissue Lipolysis and Increased Hepatic Acetyl CoA Content, PC Activity, and HGP

In order to directly examine the role of IL-6 in mediating increased HGP through increased lipolysis, hepatic acetyl CoA content, and PC activity, we performed a 4 hr intra-arterial infusion of recombinant IL-6 in control rats (Figure S6W). We found that an infusion of IL-6 caused an increase in plasma glucose, NEFA, and insulin concentrations, which were associated with a marked increase in rates of whole-body lipolysis. These increases in rates of lipolysis were associated with increases in HGP, which could be attributed to a doubling in hepatic acetyl CoA content and an 80% increase in hepatic PC activity without any change in plasma glucagon concentrations (Figures 5G–5J and S6X–S6EE). However, these effects were fully abrogated by ATGL inhibition. The induction of lipolysis in IL-6-infused rats was associated with an increase in expression of CGI-58, a potent co-activator of ATGL, in WAT. These data suggest that IL-6 stimulation of lipolysis may be attributed in part to IL-6-induced increases in WAT CGI-58 concentrations (Figure S6FF), highlighting a potential novel role for CGI-58 in inflammation-induced increases in



**Figure 5. Increases in Plasma Interleukin-6 Concentrations Cause Hyperglycemia in High-Fat-Fed Rats**

(A and B) IL-6 concentrations in plasma and white adipose tissue macrophages.

(C–F) Hepatic glucose production, PC activity, liver acetyl CoA, and whole-body palmitate turnover in high-fat-fed rats treated with an IL-6 neutralizing antibody.

(G and H) Hepatic glucose production, liver PC activity, hepatic acetyl CoA, and whole-body palmitate turnover in control rats treated with or without IL-6 or IL-6 and atglistatin (200  $\mu\text{mol/kg}$ ). In all panels, data are mean  $\pm$  SEM of  $n = 6$  per group. Comparisons by the two-tailed unpaired Student's t test (A–E) or by ANOVA with Bonferroni's multiple comparisons test (F–H). \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$  versus controls; §§ $p < 0.01$ , §§§ $p < 0.001$  versus chow-fed, IL-6-infused rats. See also Figure S6.

and WAT macrophage IL-6 concentrations (Figures 6A and 6B), are protected from high-fat diet-induced hepatic insulin resistance. We found that macrophage-specific  $m\phi$ JNK KO mice exhibited improved suppression of HGP during the hyperinsulinemic-euglycemic clamp, which was associated with reductions in clamped glycerol and palmitate turnover, hepatic acetyl CoA content, and hepatic PC activity (Figures 6C–6H and S7A–S7C). All of these effects were abrogated by infusion of IL-6 to match plasma IL-6 concentrations in the  $m\phi$ JNK knockouts to the wild-type group, demonstrating a critical role for IL-6-induced WAT lipolysis in mediating hepatic insulin resistance in high-fat-fed mice.

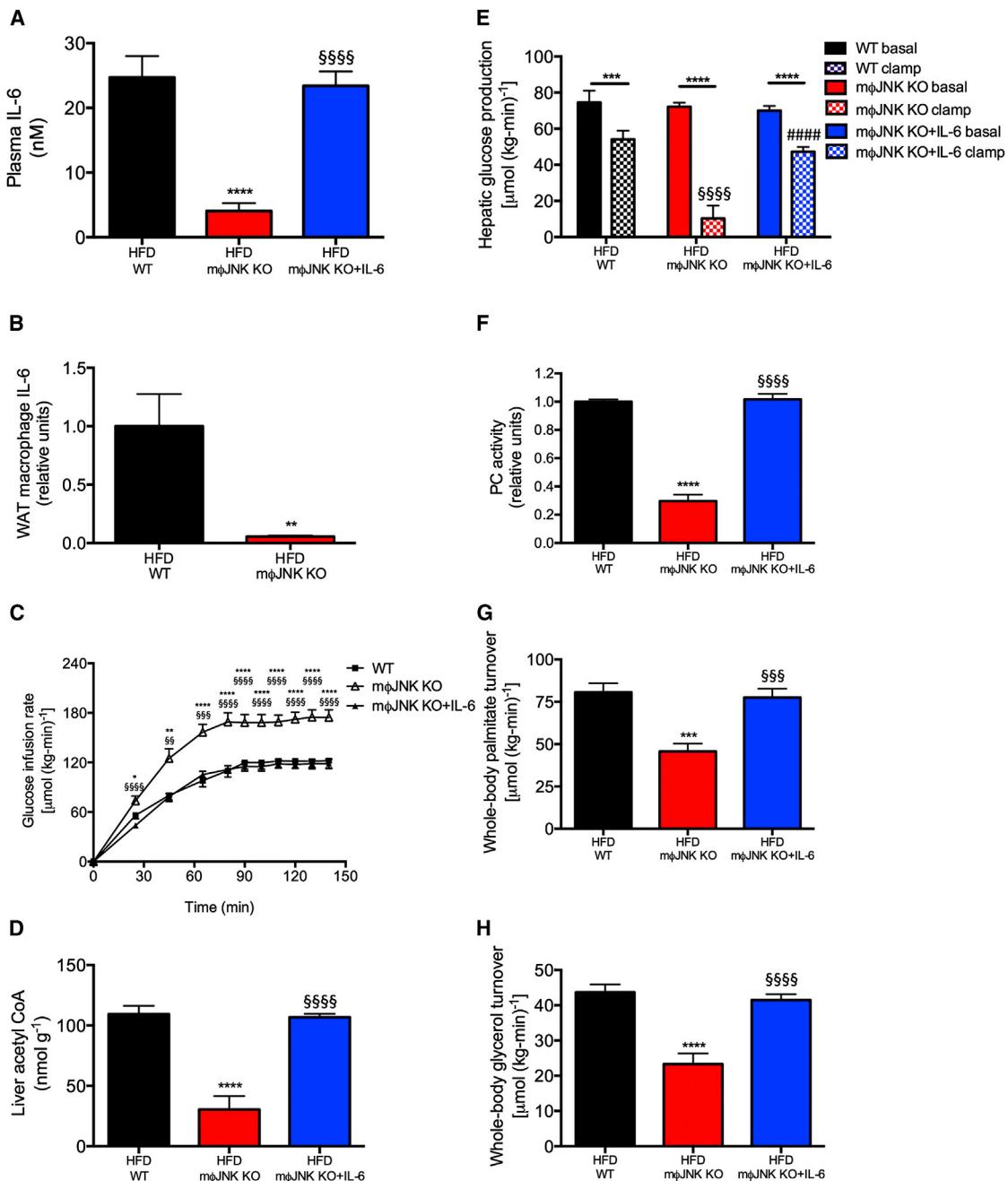
#### Obese Insulin-Resistant Adolescents Display Increased Lipolysis and Increased WAT IL-6

In order to determine whether our findings in rats would translate to humans, we measured rates of HGP and lipolysis in obese insulin-resistant adolescents. Compared to weight-matched obese insulin-sensitive control subjects, insulin-resistant subjects displayed fasting hyperglycemia and hyperinsulinemia, which were associated with increased rates of HGP, impaired insulin-mediated suppression of lipolysis and impaired insulin-mediated suppression of HGP. Impaired insulin suppression of lipolysis was associated with an increased diameter of large adipocytes and increased macrophage infiltration into the subcutaneous white adipose tissue (Figures 7A–7I; Table S1). Consistent with our findings in insulin-resistant high-fat-fed rodents, insulin-resistant obese adolescents displayed increased plasma IL-6 concentrations, and an approximately

WAT lipolysis in insulin-resistant rats. Taken together, these data are consistent with the hypothesis that local production of IL-6 from macrophages in WAT promotes increased lipolysis leading to increased hepatic acetyl CoA content, PC activity, HGP, and hyperglycemia in high-fat-fed rodents.

#### Macrophage-Specific JNK Knockout Mice Are Protected from Diet-Induced Insulin Resistance

In order to examine whether local production of IL-6, from macrophages in WAT, promotes increased lipolysis, hepatic acetyl CoA content, PC activity, HGP, and hyperglycemia in high-fat-fed rodents, we next examined whether this mechanism would explain why  $m\phi$ JNK KO mice, which have strikingly low plasma



**Figure 6. Mice Lacking JNK in Macrophages Are Protected from Diet-Induced Hepatic Insulin Resistance**

(A and B) Plasma and WAT macrophage IL-6 concentrations.

(C) Glucose infusion rate to maintain euglycemia during a hyperinsulinemic-euglycemic clamp.

(D) Liver acetyl CoA concentrations.

(E) Hepatic glucose production.

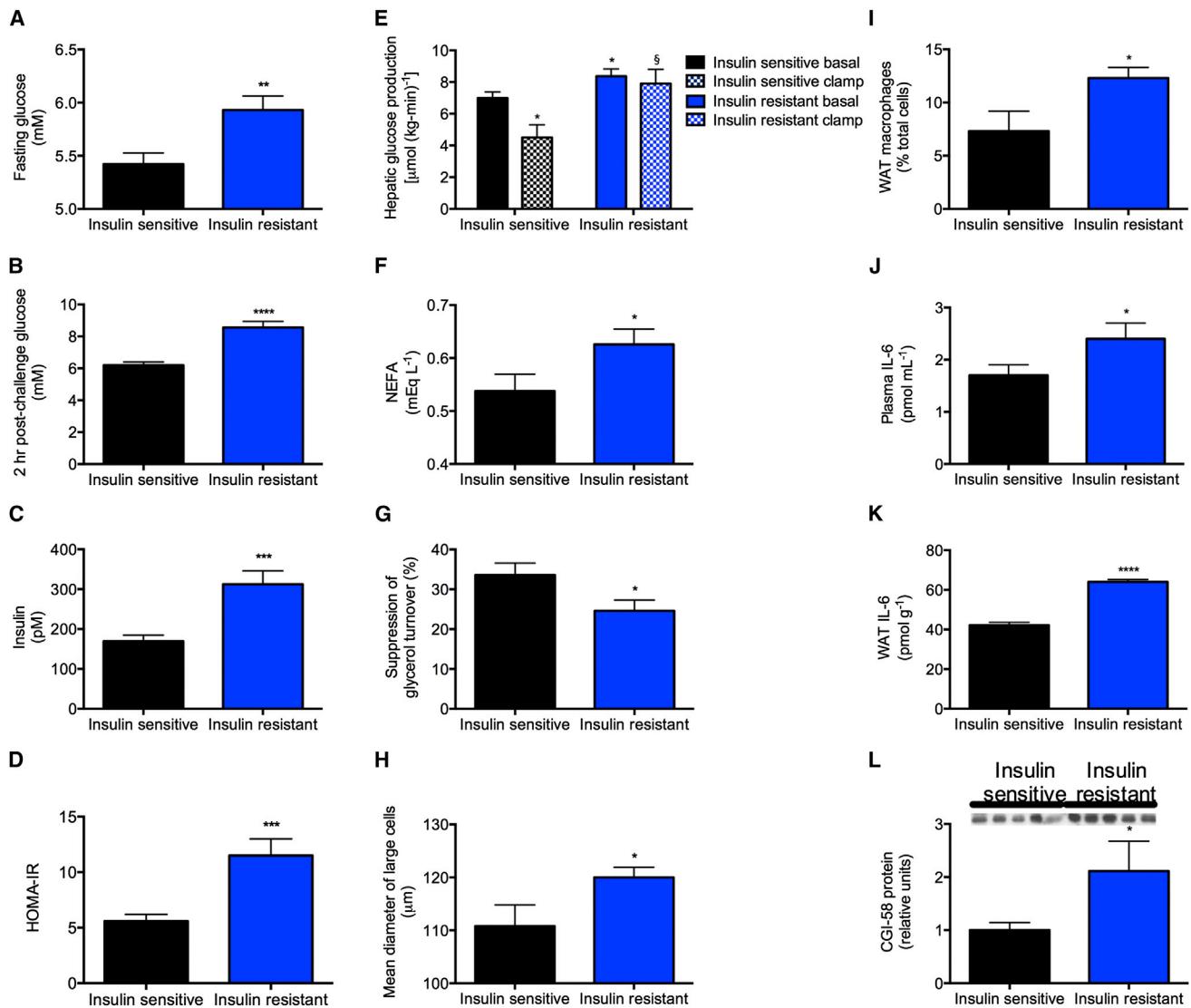
(F) Liver PC activity.

(G and H) Whole-body palmitate and glycerol turnover.

\*\*p < 0.01, \*\*\*p < 0.001, \*\*\*\*p < 0.0001 versus basal mice; §p < 0.05, §§p < 0.001, §§§p < 0.0001 versus wild-type, clamped mice. Data are mean ± SEM of n = 10 (WT), 8 (mφJNK KO), or 10 (mφJNK KO+IL-6) per group. The groups were compared by ANOVA with Bonferroni's multiple comparisons test. See also Figure S7.

50% increase in IL-6 concentrations in WAT, which were more than 20 times higher than plasma IL-6 concentrations (Figures 7J and 7K). Similar to the HFD rats, the increased rates of lipol-

ysis and HGP in obese insulin-resistant adolescents were associated with increased CGI-58 protein expression in WAT (Figure 7L).



**Figure 7. Insulin-Resistant Obese Adolescents Have Increased Lipolysis and Impaired Suppression of HGP Associated with Increased WAT IL-6 Concentrations**

(A and B) Fasting and 2-hr post-challenge plasma glucose.

(C) Fasting plasma insulin.

(D) HOMA.

(E) Hepatic glucose production. Closed bars, basal; open bars, hyperinsulinemic-euglycemic clamp.

(F) Fasting NEFA.

(G) Suppression of glycerol turnover in the clamp.

(H) Mean diameter of large adipocytes.

(I) WAT macrophage counts.

(J and K) Plasma and WAT IL-6 concentrations.

(L) WAT CGI-58 protein.

Data are the mean  $\pm$  SEM of 9–21 controls and 15–39 insulin-resistant subjects. Comparisons were made by the two-tailed unpaired Student's t test. See also Table S1.

## DISCUSSION

Suppression of HGP by insulin is critical for maintenance of normal glucose homeostasis and alterations in this process play a major role in the pathogenesis of T2D; however, the molecular mechanism by which insulin suppresses HGP has re-

mained elusive and has been attributed to direct actions of insulin on liver (Edgerton et al., 2009; Hepp, 1977; Pilkins and Granner, 1992) and indirect effects of insulin on brain (Hill et al., 2010; Jordan et al., 2010; Lin et al., 2010; Obici et al., 2002; Schwartz et al., 2013) and adipocyte (Lewis et al., 1996; Rebrin et al., 1996). The canonical signaling pathway by which insulin has been

postulated to directly inhibit hepatic gluconeogenesis is through activation of Akt, resulting in suppression of gluconeogenic protein expression by phosphorylation and exclusion of Forkhead box (Foxo1) from the nucleus of the hepatocyte, and consequent transcription-mediated reduction in hepatic gluconeogenesis (Haeusler et al., 2014). However, the finding that triple knockout mice lacking Akt 1 and Akt2 and Foxo1 in liver have normal suppression of HGP during a hyperinsulinemic-euglycemic clamp (Lu et al., 2012) poses considerable challenges to this paradigm. Furthermore, insulin suppresses HGP within minutes of administration *in vivo*, which can unlikely be explained by transcriptionally mediated processes.

To explain this apparent paradox, we hypothesized that insulin's ability to suppress lipolysis in WAT may be critical for the suppression of HGP by reducing fatty acid flux to the liver resulting in decreased hepatic acetyl CoA concentrations and decreased PC activity resulting in decreased conversion of pyruvate to glucose. While hepatic acetyl CoA has been long known to modulate pyruvate carboxylase activity *in vitro* (Barritt et al., 1966; Cazzullo and Stoppani, 1968; Cooper and Benedict, 1966; Keech and Utter, 1963; Krebs et al., 1965), methodological limitations stemming from acetyl CoA's low hepatocellular concentrations and rapid degradation *ex vivo* have prevented its measurement *in vivo*. Additionally, Chisholm et al. were not able to demonstrate any direct effect of insulin on pyruvate carboxylase activity in isolated hepatocytes (Chisholm et al., 1983), begging the question of whether a hepatocyte-autonomous process is primarily responsible for insulin-mediated suppression of PC activity and hepatic glucose production.

In order to examine this hypothesis, we developed a comprehensive metabolomics flux approach to simultaneously measure whole-body rates of lipolysis, glucose turnover, and intrahepatic fluxes as well as hepatic acetyl CoA content in rapidly freeze-clamped livers. Using this approach, we demonstrate that hepatic acetyl CoA content is a key regulator of HGP and that an insulin-mediated reduction of hepatic acetyl CoA concentrations, through inhibition of WAT lipolysis, is the major mechanism by which insulin suppresses HGP *in vivo*. While the ability of insulin to suppress HGP through an indirect mechanism was postulated by Levine and Fritz more than 50 years ago, the molecular mechanism regarding this effect has remained elusive. These are the first studies, to our knowledge, that demonstrate insulin regulation of pyruvate carboxylase activity through modulation of hepatic acetyl CoA concentrations and demonstrate that this mechanism is responsible for insulin suppression of hepatic glucose production *in vivo*. Furthermore, these data identify the mechanism by which mice lacking the canonical hepatic insulin signaling factors (Akt1, Akt2, Foxo1) are able to suppress HGP normally and explain the enigmatic inability of insulin to suppress hepatic glucose production in isolated hepatocyte preparations (Lu et al., 2012).

We next examined whether increased hepatic acetyl CoA concentrations might be responsible for increased rates of hepatic glucose production in insulin resistance and might link inflammation and macrophage-induced lipolysis in WAT with fasting and postprandial hyperglycemia through a similar mechanism. While lipid-induced insulin resistance occurs early in the pathogenesis

of T2D and can be dissociated from inflammation at this stage, a key question arises as to identification of the factor(s) that promotes the progression from ectopic lipid-induced insulin resistance to impaired glucose tolerance and fasting hyperglycemia (Shulman, 2014). The classic view of this process implicates impaired pancreatic  $\beta$ -cell and  $\alpha$ -cell function along with inflammation in this transition where the former leads to increased hepatic gluconeogenic gene transcription and the latter inhibits insulin action through the release of cytokines and adipocytokines, which, in turn, leads to inhibition of insulin signaling and increased hepatic gluconeogenic protein expression through activation of the NF- $\kappa$ B/JNK/ceramide biosynthetic pathways (Olefsky and Glass, 2010). In this study, we explored an alternative hypothesis in which progression to fasting hyperglycemia occurs through macrophage-induced WAT lipolysis resulting in increased hepatic pyruvate carboxylase activity through increases in hepatic acetyl CoA content and increased conversion of glycerol to glucose. Consistent with this hypothesis, we show that high-fat-fed rats manifested a 2-fold increase in plasma and WAT IL-6 concentrations, which caused hyperglycemia by increasing lipolysis, hepatic acetyl CoA, and HGP both before and during a hyperinsulinemic-euglycemic clamp. Infusions of IL-6 and treatment of high-fat-fed rats with an IL-6-neutralizing antibody demonstrate that IL-6 is both necessary and sufficient to induce hyperglycemia in the rat. These findings were validated genetically in adipose-specific ATGL knockout mice and m $\phi$ JNK knockout mice, which were protected from high-fat feeding-induced hepatic insulin resistance due to reductions in WAT lipolysis leading to decreased hepatic acetyl CoA content, decreased PC activity, and decreased rates of HGP during a hyperinsulinemic-euglycemic clamp.

Finally, we show these results translate to humans by demonstrating that WAT inflammation is associated with increased IL-6 concentrations in WAT, increased rates of lipolysis, increased rates of HGP, and hepatic insulin resistance assessed by a hyperinsulinemic-euglycemic clamp in obese insulin-resistant adolescents. These results are consistent with previous studies, which have observed that increased plasma IL-6 concentrations are strong predictors of insulin resistance in humans (Fernandez-Real et al., 2001; Glund et al., 2007; Kado et al., 1999; Mohamed-Ali et al., 1997; Petersen et al., 2006; Pickup et al., 1997; Steptoe et al., 2014; Straub et al., 2000; Yuen et al., 2009).

Taken together, these data demonstrate that allosteric regulation of pyruvate carboxylase activity by hepatic acetyl CoA, through insulin suppression of WAT lipolysis, is a critical factor in the regulation of hepatic glucose production by insulin *in vivo*. Furthermore, these studies demonstrate that macrophage-induced WAT lipolysis leading to increased hepatic acetyl CoA content and increased pyruvate carboxylase activity/flux is a key molecular mechanism linking WAT inflammation to both fasting and postprandial hyperglycemia in T2D. These data explain insulin's inability to suppress hepatic glucose production in high-fat-fed rodent models of diet-induced obesity and hepatic insulin resistance independent of transcriptional regulation of gluconeogenic enzymes or alterations in hepatic glycogen metabolism and also explain recent studies that find no relationship between hepatic gluconeogenic protein

expression and fasting hyperglycemia in obese humans (Kumashiro et al., 2011). These data also challenge the canonical view of inflammation-mediated hepatic insulin resistance through alterations in circulating cytokines, which, in turn, lead to inhibition of hepatic insulin signaling and increased hepatic gluconeogenic protein transcription through activation of the NF- $\kappa$ B/JNK/ceramide biosynthetic pathways. Finally, these results have important clinical implications for the pathogenesis of T2D by linking inflammation and macrophage-induced WAT lipolysis leading to increased hepatic acetyl CoA concentrations and increased PC flux resulting in fasting and postprandial hyperglycemia and identify macrophage-induced WAT lipolysis and hepatic acetyl CoA as a novel potential therapeutic target for T2D.

## EXPERIMENTAL PROCEDURES

### Animal Studies

Rodent protocols were approved by the International Animal Care and Use Committees at the institutions at which the animals were housed (Yale University, the University of Pennsylvania, and the University of Massachusetts). Additional information about the animal models is provided in [Extended Experimental Procedures](#).

### Basal and Hyperinsulinemic-Euglycemic Clamp Studies

Basal infusion and hyperinsulinemic-euglycemic clamp studies were used to assess whole-body and hepatic insulin sensitivity as described in the [Extended Experimental Procedures](#).

### Acute Suppression of Hepatic Gluconeogenesis

To assess insulin's acute ability to suppress gluconeogenesis, chow-fed rats were infused with [ $^3$ H] glucose (0.1  $\mu$ Ci/min) for 120 min. Blood samples were taken at 100, 110, and 120 min to assess basal glucose turnover. At 120 min, a 20-min insulin infusion was initiated ( $20 \text{ mU} \times \text{kg}^{-1}$  prime,  $2 \text{ mU} \times \text{kg}^{-1} \times \text{min}^{-1}$  for the duration of the study). Blood samples (100  $\mu$ l) were taken every 3 min during the insulin infusion, and specific activity was measured using a scintillation counter.

### Variable Glycerol and Acetate Infusions

Sodium acetate was infused intra-arterially in rats at infusion rates 0, 130, 195, 260, or 390  $\mu\text{mol} \times \text{kg}^{-1} \times \text{min}^{-1}$ . [ $^3$ H] glucose was co-infused and blood and livers obtained as above. In a separate group of rats, [ $1,1,2,3,3-\text{d}_5$ ] glycerol was infused at infusion rates 0, 10, 30, 50, or 70  $\mu\text{mol} \times \text{kg}^{-1} \times \text{min}^{-1}$  for 120 min. Glycerol turnover and concentrations were measured by gas chromatography/mass spectrometry (GC/MS) as described below. In the studies in which acetate and glycerol were co-infused during the clamp to abrogate insulin-mediated suppression of hepatic glucose production, the infusion rates selected were  $195 \mu\text{mol} \times \text{kg}^{-1} \times \text{min}^{-1}$  acetate and  $10 \mu\text{mol} \times \text{kg}^{-1} \times \text{min}^{-1}$  glycerol.

### Measurement of Glucose, Fatty Acid, and Glycerol Turnover

Whole-body glucose turnover was calculated as the ratio of the [ $^3$ H] glucose infusion rate to the specific activity of plasma glucose at the end of the 120 min infusion. Endogenous glucose production represents the difference between the glucose infusion rate and the rate of glucose appearance. The liver was assumed to contribute 90% of whole-body glucose production based on our prior studies (Perry et al., 2014). Hepatic flux rates were calculated as we describe in [Extended Experimental Procedures](#). [ $^1\text{U}-^{13}\text{C}$ ] palmitate (Beylot et al., 1993) and [ $1,1,2,3,3-\text{d}_5$ ] glycerol (Previs et al., 1999) enrichment and total glycerol concentration were measured by GC/MS as previously reported, and turnover was calculated using the formula

$$\text{Turnover} = \left( \frac{\text{tracer enrichment}}{\text{plasma enrichment}} - 1 \right) * \text{Infusion rate.}$$

### ATGL Inhibition

Groups of overnight fasted chow and high-fat-fed rats were treated with intra-peritoneal injections of atglistatin (200 or 600  $\mu\text{mol}/\text{kg}$ , as designated in the text, XcessBiosciences), a small-molecule ATGL inhibitor, or saline vehicle. They were then subjected to a basal infusion of [ $1,1,2,3,3-\text{d}_5$ ] glycerol, [ $^1\text{U}-^{13}\text{C}$ ] palmitate, and [ $^3$ H] glucose as described above.

### CPT-1 Inhibition

Overnight fasted chow and high-fat-fed rats were treated with intraperitoneal injections of etomoxir (8 mg  $\times \text{kg}^{-1}$ ) and immediately began an infusion of [ $1,1,2,3,3-\text{d}_5$ ] glycerol, [ $^1\text{U}-^{13}\text{C}$ ] palmitate, and [ $^3$ H] glucose as described above.

### Adipocyte Morphology and Function

White adipose tissue glucose uptake was measured in basal and clamped rats following injection of 2-[ $1-^{14}\text{C}$ ] deoxyglucose (Cartee and Bohn, 1995). Insulin-stimulated WAT lipogenesis and adipocyte cell size were measured in chow and 4-week high-fat-fed rats as previously described (Kursawe et al., 2010). To measure IL-6-induced lipolysis, ~100-mg WAT samples were incubated in 5 mM glucose media containing 15 or 25 nM IL-6 to mimic plasma IL-6 concentrations in vivo in chow and high-fat-fed rats. After 6 hr of incubation at 37°C, NEFA concentrations in the media were measured and normalized to the total protein content of the WAT sample, measured by Bradford assay.

### IL-6 Neutralization

Overnight fasted 4-week high-fat-fed rats were treated with an intraperitoneal injection of a neutralizing antibody to IL-6 (10  $\mu\text{g}/\text{rat}$ , R&D Systems) in saline, or an equivalent volume of saline. Immediately following the antibody injection, [ $1,1,2,3,3-\text{d}_5$ ] glycerol, [ $^1\text{U}-^{13}\text{C}$ ] palmitate, and [ $^3$ H] glucose were infused intra-arterially for the measurement of lipolysis and hepatic glucose production as described above.

### IL-6 Infusion

A continuous intra-arterial infusion of IL-6 (10  $\mu\text{g}/\text{hr}^{-1}$ , Sigma) was performed for 4 hr. After 2 hr of IL-6 infusion, atglistatin (200  $\mu\text{mol}/\text{kg}^{-1}$ ) was injected intravenously, and rates of lipolysis and hepatic glucose production were assessed as described above.

### Biochemical Analyses

Plasma glucose concentrations were measured using the YSI 2700 Select Biochemistry Analyzer, and plasma lactate concentrations by COBAS. Plasma insulin concentrations were measured by ELISA (Millipore). NEFA concentrations were measured using a Wako kit. Plasma insulin was measured by radioimmunoassay by the Yale Diabetes Research Core. Plasma FGF-21, adiponectin, corticosterone, epinephrine, norepinephrine, growth hormone, and inflammatory cytokine concentrations were measured by ELISA (Millipore, Invitrogen, Abcam, IBL International, IBL International, Life Technologies, and QIAGEN, respectively).

### Tissue Analyses

Liver acetyl CoA concentration and enrichment was measured by liquid chromatography/mass spectrometry/mass spectrometry (LC-MS/MS) as described in the [Extended Experimental Procedures](#). Liver glycogen concentrations were measured in basal and clamped rats using the method of Dalrymple and Hamm (1973). Pyruvate carboxylase activity was measured as previously described (Vatner et al., 2013), with acetyl CoA concentrations adjusted to match those measured by LC-MS/MS in each group. Liver ATP, ADP, AMP, NAD, and NADH concentrations were measured by LC-MS/MS as previously described (Perry et al., 2013).

### WAT Macrophage Isolation

WAT tissue was washed with HBSS and then incubated at 37°C, shaking for 60 min in HBSS with 0.5 mg/ml collagenase (Sigma). After washing three times with HBSS, cell debris was reduced by Histopaque density centrifugation. The remaining cells were stained with CD11b-PE (eBioscience), and CD11b+ WAT macrophages were enriched to >90% purity by EasySep PE

positive selection kit (STEMCELL Technologies). Purity and cell number were confirmed by flow cytometry. WAT macrophage IL-6 and WAT tissue GM-CSF and TNF- $\alpha$  were measured by ELISA (Life Technologies, Abcam, and SABiosciences, respectively) and normalized to total protein content measured by Bradford assay.

#### Western Blots

Protein expression was measured in mouse and rat liver by western blot, with antibodies obtained from Cell Signaling Technology (Akt1, Akt2, Akt pSer473, Foxo1, pyruvate carboxylase, phosphoenolpyruvate carboxykinase, glucose-6-phosphatase, adipose triglyceride lipase, insulin receptor substrate-1, and glyceraldehyde 3-phosphate dehydrogenase) or from Santa Cruz Biotechnology (CGI-58, peroxisome proliferator-activated receptor gamma, glucose transporter type 4, and protein phosphatase 1).

#### Human Studies

The Yale Pathophysiology of Type 2 Diabetes in Obese Youth Study is a long-term project aimed at examining early alterations in glucose metabolism in relation to fat patterning in obese adolescents. Specific details of this study are described in the [Extended Experimental Procedures](#).

#### Statistical Analyses

Statistical analyses were performed using GraphPad Prism. Comparisons between two groups were performed by the two-tailed unpaired Student's t test, while three groups were compared by ANOVA with Bonferroni's post hoc test. Differences were considered significant at  $p < 0.05$ . Data are presented as mean  $\pm$  SEM.

#### SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, seven figures, and one table and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2015.01.012>.

#### AUTHOR CONTRIBUTIONS

R.J.P. and G.I.S. designed the project and the experimental protocols. P.M.T. and M.J.B. generated the Akt1/2/Foxo1 knockout mice, and H.-B.R. and X.Y. performed adenoviral cre injections to ablate the proteins of interest. M.S.H. and R.J.D. generated the m $\phi$ JNK knockout mice, and H.S.S. generated the adipose-specific ATGL-knockout mice. R.J.P. performed the *in vivo* rat studies, J.-P.G.C., M.J.J., and A.A. performed the *in vivo* mouse studies, and R.K. and S.C. performed the *in vivo* human studies. R.J.P., J.-P.G.C., D.Z., R.K., C.J.P., M.J.J., S.M.K., and G.W.C. contributed to plasma and tissue analysis. K.F.P. contributed to important discussions of the protocols and data. The manuscript was written by R.J.P. and G.I.S. with input from all other authors.

#### ACKNOWLEDGMENTS

The authors thank Dr. James Dziura for assistance with the statistical analysis, Isis Pharmaceuticals and Dr. Sanjay Bhanot for providing us with the 2'-O-methoxyethyl chimeric antisense oligonucleotide for the insulin receptor kinase, and Dr. Jianying Dong, Mario Kahn, John Stack, Gina Butrico, Ali Nasiri, Maria Batsu, and Wanling Zhu for their skilled technical assistance. This study was funded by grants from the NIH (R01 DK-40936, R24 DK-085638, R01 AG-023686, P30 DK-45735, U24 DK-059635, T32 DK-101019, R01 DK-056886, R01 DK-093959, R01 NS-087568, R01 DK93928, UL1 TR-000142, R01-HD028016, R01-HD 04787, R01 DK085577, R24 DK-090963) and the Novo Nordisk Foundation Center for Basic Metabolic Research.

Received: August 3, 2014

Revised: November 10, 2014

Accepted: January 5, 2015

Published: February 5, 2015

#### REFERENCES

Barritt, G.J., Keech, D.B., and Ling, A.M. (1966). Apparent co-operative effect of acetyl-CoA on sheep kidney pyruvate carboxylase. *Biochem. Biophys. Res. Commun.* 24, 476–481.

Beylot, M., David, F., and Brunengraber, H. (1993). Determination of the 13C-labeling pattern of glutamate by gas chromatography-mass spectrometry. *Anal. Biochem.* 212, 532–536.

Boyle, J.P., Thompson, T.J., Gregg, E.W., Barker, L.E., and Williamson, D.F. (2010). Projection of the year 2050 burden of diabetes in the US adult population: dynamic modeling of incidence, mortality, and prediabetes prevalence. *Popul. Health Metr.* 8, 29.

Cartee, G.D., and Bohn, E.E. (1995). Growth hormone reduces glucose transport but not GLUT-1 or GLUT-4 in adult and old rats. *Am. J. Physiol.* 268, E902–E909.

Cazzulo, J.J., and Stoppani, A.O. (1968). The regulation of yeast pyruvate carboxylase by acetyl-coenzyme A and L-aspartate. *Arch. Biochem. Biophys.* 127, 563–567.

Cernkovich, E.R., Deng, J., Bond, M.C., Combs, T.P., and Harp, J.B. (2008). Adipose-specific disruption of signal transducer and activator of transcription 3 increases body weight and adiposity. *Endocrinology* 149, 1581–1590.

Chisholm, A.B., Allan, E.H., and Titheradge, M.A. (1983). Regulation of mitochondrial pyruvate carboxylation in isolated hepatocytes by acute insulin treatment. *Biochem. J.* 214, 451–458.

Cooper, T.G., and Benedict, C.R. (1966). The participation of acetyl-CoA in pyruvate carboxylase. *Biochem. Biophys. Res. Commun.* 22, 285–290.

Dalrymple, R.H., and Hamm, D. (1973). A method for the extraction of glycogen and metabolites from a single muscle sample. *Int. J. Food Sci. Technol.* 8, 439–444.

Edgerton, D.S., Ramnanan, C.J., Grueter, C.A., Johnson, K.M., Lautz, M., Neal, D.W., Williams, P.E., and Cherrington, A.D. (2009). Effects of insulin on the metabolic control of hepatic gluconeogenesis *in vivo*. *Diabetes* 58, 2766–2775.

Fernandez-Real, J.M., Vayreda, M., Richart, C., Gutierrez, C., Broch, M., Vendrell, J., and Ricart, W. (2001). Circulating interleukin 6 levels, blood pressure, and insulin sensitivity in apparently healthy men and women. *J. Clin. Endocrinol. Metab.* 86, 1154–1159.

Fonseca-Alaniz, M.H., Takada, J., Alonso-Vale, M.I., and Lima, F.B. (2007). Adipose tissue as an endocrine organ: from theory to practice. *J. Pediatr. (Rio J.)* 83, S192–S203.

Glund, S., Deshmukh, A., Long, Y.C., Moller, T., Koistinen, H.A., Caidahl, K., Zierath, J.R., and Krook, A. (2007). Interleukin-6 directly increases glucose metabolism in resting human skeletal muscle. *Diabetes* 56, 1630–1637.

Guschin, D., Rogers, N., Briscoe, J., Witthuhn, B., Watling, D., Horn, F., Pellegrini, S., Yasukawa, K., Heinrich, P., Stark, G.R., et al. (1995). A major role for the protein tyrosine kinase JAK1 in the JAK/STAT signal transduction pathway in response to interleukin-6. *EMBO J.* 14, 1421–1429.

Haeusler, R.A., Hartil, K., Vaitheeswaran, B., Arrieta-Cruz, I., Knight, C.M., Cook, J.R., Kammoun, H.L., Febbraio, M.A., Gutierrez-Juarez, R., Kurland, I.J., and Accili, D. (2014). Integrated control of hepatic lipogenesis versus glucose production requires FoxO transcription factors. *Nat. Commun.* 5, 5190.

Han, M.S., Jung, D.Y., Morel, C., Lakhani, S.A., Kim, J.K., Flavell, R.A., and Davis, R.J. (2013). JNK expression by macrophages promotes obesity-induced insulin resistance and inflammation. *Science* 339, 218–222.

Hepp, K.D. (1977). Studies on the mechanism of insulin action: basic concepts and clinical implications. *Diabetologia* 13, 177–186.

Hill, J.W., Elias, C.F., Fukuda, M., Williams, K.W., Berglund, E.D., Holland, W.L., Cho, Y.R., Chuang, J.C., Xu, Y., Choi, M., et al. (2010). Direct insulin and leptin action on pro-opiomelanocortin neurons is required for normal glucose homeostasis and fertility. *Cell Metab.* 11, 286–297.

Jordan, S.D., Könner, A.C., and Brüning, J.C. (2010). Sensing the fuels: glucose and lipid signaling in the CNS controlling energy homeostasis. *Cell. Mol. Life Sci.* 67, 3255–3273.

Kado, S., Nagase, T., and Nagata, N. (1999). Circulating levels of interleukin-6, its soluble receptor and interleukin-6/interleukin-6 receptor complexes in patients with type 2 diabetes mellitus. *Acta Diabetol.* 36, 67–72.

Kahn, S.E., Cooper, M.E., and Del Prato, S. (2014). Pathophysiology and treatment of type 2 diabetes: perspectives on the past, present, and future. *Lancet* 383, 1068–1083.

Keech, D.B., and Utter, M.F. (1963). Pyruvate Carboxylase. II. Properties. *J. Biol. Chem.* 238, 2609–2614.

Krebs, H.A., Speak, R.N., and Hems, R. (1965). Acceleration of Renal Gluconeogenesis by Ketone Bodies and Fatty Acids. *Biochem. J.* 94, 712–720.

Kumashiro, N., Erion, D.M., Zhang, D., Kahn, M., Beddow, S.A., Chu, X., Still, C.D., Gerhard, G.S., Han, X., Dziura, J., et al. (2011). Cellular mechanism of insulin resistance in nonalcoholic fatty liver disease. *Proc. Natl. Acad. Sci. USA* 108, 16381–16385.

Kursawe, R., Eszlinger, M., Narayan, D., Liu, T., Bazuine, M., Cali, A.M., D'Adamo, E., Shaw, M., Pierpont, B., Shulman, G.I., et al. (2010). Cellularity and adipogenic profile of the abdominal subcutaneous adipose tissue from obese adolescents: association with insulin resistance and hepatic steatosis. *Diabetes* 59, 2288–2296.

Lewis, G.F., Zinman, B., Groenewoud, Y., Vranic, M., and Giacca, A. (1996). Hepatic glucose production is regulated both by direct hepatic and extrahepatic effects of insulin in humans. *Diabetes* 45, 454–462.

Lin, H.V., Plum, L., Ono, H., Gutiérrez-Juárez, R., Shanabrough, M., Borok, E., Horvath, T.L., Rossetti, L., and Accili, D. (2010). Divergent regulation of energy expenditure and hepatic glucose production by insulin receptor in agouti-related protein and POMC neurons. *Diabetes* 59, 337–346.

Lu, M., Wan, M., Leavens, K.F., Chu, Q., Monks, B.R., Fernandez, S., Ahima, R.S., Ueki, K., Kahn, C.R., and Birnbaum, M.J. (2012). Insulin regulates liver metabolism in vivo in the absence of hepatic Akt and Foxo1. *Nat. Med.* 18, 388–395.

Lütticken, C., Wegenka, U.M., Yuan, J., Buschmann, J., Schindler, C., Ziemiecki, A., Harpur, A.G., Wilks, A.F., Yasukawa, K., Taga, T., et al. (1994). Association of transcription factor APRF and protein kinase Jak1 with the interleukin-6 signal transducer gp130. *Science* 263, 89–92.

Manteiga, S., Choi, K., Jayaraman, A., and Lee, K. (2013). Systems biology of adipose tissue metabolism: regulation of growth, signaling and inflammation. *Wiley Interdiscip Rev Syst Biol Med* 5, 425–447.

Mohamed-Ali, V., Goodrick, S., Rawesh, A., Katz, D.R., Miles, J.M., Yudkin, J.S., Klein, S., and Coppack, S.W. (1997). Subcutaneous adipose tissue releases interleukin-6, but not tumor necrosis factor-alpha, in vivo. *J. Clin. Endocrinol. Metab.* 82, 4196–4200.

Obici, S., Zhang, B.B., Karkanias, G., and Rossetti, L. (2002). Hypothalamic insulin signaling is required for inhibition of glucose production. *Nat. Med.* 8, 1376–1382.

Olefsky, J.M., and Glass, C.K. (2010). Macrophages, inflammation, and insulin resistance. *Annu. Rev. Physiol.* 72, 219–246.

Perry, R.J., Kim, T., Zhang, X.M., Lee, H.Y., Pesta, D., Popov, V.B., Zhang, D., Rahimi, Y., Jurczak, M.J., Cline, G.W., et al. (2013). Reversal of hypertriglyceridemia, fatty liver disease, and insulin resistance by a liver-targeted mitochondrial uncoupler. *Cell Metab.* 18, 740–748.

Perry, R.J., Zhang, X.M., Zhang, D., Kumashiro, N., Camporez, J.P., Cline, G.W., Rothman, D.L., and Shulman, G.I. (2014). Leptin reverses diabetes by suppression of the hypothalamic-pituitary-adrenal axis. *Nat. Med.* 20, 759–763.

Petersen, K.F., Dufour, S., Feng, J., Befroy, D., Dziura, J., Dalla Man, C., Cobelli, C., and Shulman, G.I. (2006). Increased prevalence of insulin resistance and nonalcoholic fatty liver disease in Asian-Indian men. *Proc. Natl. Acad. Sci. USA* 103, 18273–18277.

Pickup, J.C., Mattock, M.B., Chusney, G.D., and Burt, D. (1997). NIDDM as a disease of the innate immune system: association of acute-phase reactants and interleukin-6 with metabolic syndrome X. *Diabetologia* 40, 1286–1292.

Pilkis, S.J., and Granner, D.K. (1992). Molecular physiology of the regulation of hepatic gluconeogenesis and glycolysis. *Annu. Rev. Physiol.* 54, 885–909.

Previs, S.F., Cline, G.W., and Shulman, G.I. (1999). A critical evaluation of mass isotopomer distribution analysis of gluconeogenesis in vivo. *Am. J. Physiol.* 277, E154–E160.

Rebrin, K., Steil, G.M., Mittelman, S.D., and Bergman, R.N. (1996). Causal linkage between insulin suppression of lipolysis and suppression of liver glucose output in dogs. *J. Clin. Invest.* 98, 741–749.

Robbins, G.R., Wen, H., and Ting, J.P. (2014). Inflammasomes and metabolic disorders: old genes in modern diseases. *Mol. Cell* 54, 297–308.

Samuel, V.T., and Shulman, G.I. (2012). Mechanisms for insulin resistance: common threads and missing links. *Cell* 148, 852–871.

Savage, D.B., Petersen, K.F., and Shulman, G.I. (2005). Mechanisms of insulin resistance in humans and possible links with inflammation. *Hypertension* 45, 828–833.

Schwartz, M.W., Seeley, R.J., Tschoep, M.H., Woods, S.C., Morton, G.J., Myers, M.G., and D'Alessio, D. (2013). Cooperation between brain and islet in glucose homeostasis and diabetes. *Nature* 503, 59–66.

Shulman, G.I. (2014). Ectopic fat in insulin resistance, dyslipidemia, and cardiometabolic disease. *N. Engl. J. Med.* 371, 1131–1141.

Stahl, N., Boulton, T.G., Farruggella, T., Ip, N.Y., Davis, S., Witthuhn, B.A., Quelle, F.W., Silvennoinen, O., Barbieri, G., Pellegrini, S., et al. (1994). Association and activation of Jak-Tyk kinases by CNTF-LIF-OSM-IL-6 beta receptor components. *Science* 263, 92–95.

Steele, R. (1959). Influences of glucose loading and of injected insulin on hepatic glucose output. *Ann. N Y Acad. Sci.* 82, 420–430.

Steptoe, A., Hackett, R.A., Lazzarino, A.I., Bostock, S., La Marca, R., Carvalho, L.A., and Hamer, M. (2014). Disruption of multisystem responses to stress in type 2 diabetes: investigating the dynamics of allostatic load. *Proc. Natl. Acad. Sci. USA* 111, 15693–15698.

Straub, R.H., Hense, H.W., Andus, T., Schölmerich, J., Riegger, G.A., and Schunkert, H. (2000). Hormone replacement therapy and interrelation between serum interleukin-6 and body mass index in postmenopausal women: a population-based study. *J. Clin. Endocrinol. Metab.* 85, 1340–1344.

Suganami, T., and Ogawa, Y. (2010). Adipose tissue macrophages: their role in adipose tissue remodeling. *J. Leukoc. Biol.* 88, 33–39.

Vatner, D.F., Weismann, D., Beddow, S.A., Kumashiro, N., Erion, D.M., Liao, X.H., Grover, G.J., Webb, P., Phillips, K.J., Weiss, R.E., et al. (2013). Thyroid hormone receptor-β agonists prevent hepatic steatosis in fat-fed rats but impair insulin sensitivity via discrete pathways. *Am. J. Physiol. Endocrinol. Metab.* 305, E89–E100.

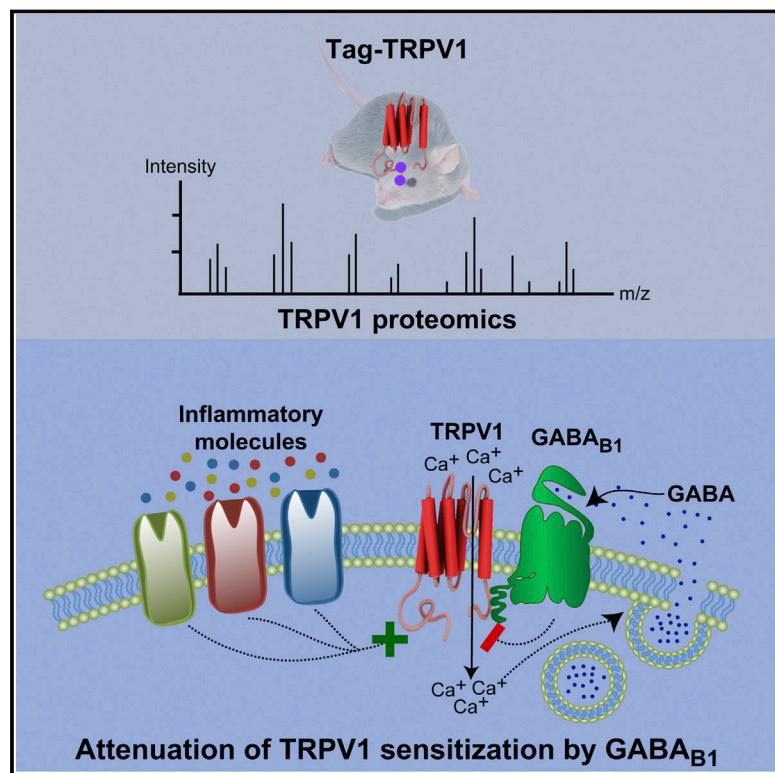
Yamaguchi, T., Omatsu, N., Matsushita, S., and Osumi, T. (2004). CGI-58 interacts with perilipin and is localized to lipid droplets. Possible involvement of CGI-58 mislocalization in Chanarin-Dorfman syndrome. *J. Biol. Chem.* 279, 30490–30497.

Yamaguchi, T., Omatsu, N., Morimoto, E., Nakashima, H., Ueno, K., Tanaka, T., Satouchi, K., Hirose, F., and Osumi, T. (2007). CGI-58 facilitates lipolysis on lipid droplets but is not involved in the vesiculation of lipid droplets caused by hormonal stimulation. *J. Lipid Res.* 48, 1078–1089.

Yuen, D.Y., Dwyer, R.M., Matthews, V.B., Zhang, L., Drew, B.G., Neill, B., Kingwell, B.A., Clark, M.G., Rattigan, S., and Febbraio, M.A. (2009). Interleukin-6 attenuates insulin-mediated increases in endothelial cell signaling but augments skeletal muscle insulin action via differential effects on tumor necrosis factor-alpha expression. *Diabetes* 58, 1086–1095.

# GABA Blocks Pathological but Not Acute TRPV1 Pain Signals

## Graphical Abstract



## Authors

Christina Hanack, Mirko Moroni, ...,  
Matthias Selbach, Jan Siemens

## Correspondence

[jan.siemens@pharma.uni-heidelberg.de](mailto:jan.siemens@pharma.uni-heidelberg.de)

## In Brief

The neurotransmitter GABA is released upon stimulation of the pain receptor TRPV1 and engages a non-canonical signaling pathway that inhibits only hyperactive TRPV1, leaving homeostatic pain responses intact.

## Highlights

- GABA<sub>B1</sub> forms a complex with TRPV1 to counteract inflammatory pain
- GABA<sub>B1</sub> modulates TRPV1 via a non-canonical, GABA<sub>B2</sub>-independent pathway
- TRPV1 activation triggers GABA release from peripheral nerve endings
- GABA serves as a modulator of nociceptor sensitization in the periphery

# GABA Blocks Pathological but Not Acute TRPV1 Pain Signals

Christina Hanack,<sup>1,2,5</sup> Mirko Moroni,<sup>2,5</sup> Wanessa C. Lima,<sup>1,5</sup> Hagen Wende,<sup>1,5</sup> Marieluise Kirchner,<sup>2</sup> Lisa Adelfinger,<sup>3</sup> Katrin Schrenk-Siemens,<sup>1</sup> Anke Tappe-Theodor,<sup>1</sup> Christiane Wetzel,<sup>2</sup> P. Henning Kuich,<sup>2</sup> Martin Gassmann,<sup>3</sup> Dennis Roggenkamp,<sup>4</sup> Bernhard Bettler,<sup>3</sup> Gary R. Lewin,<sup>2</sup> Matthias Selbach,<sup>2</sup> and Jan Siemens<sup>1,\*</sup>

<sup>1</sup>Department of Pharmacology, University of Heidelberg, Im Neuenheimer Feld 366, 69120 Heidelberg, Germany

<sup>2</sup>Max Delbrück Center for Molecular Medicine, Robert-Rössle-Strasse 10, 13125 Berlin, Germany

<sup>3</sup>Department of Biomedicine, University of Basel, CH-4056 Basel, Switzerland

<sup>4</sup>Beiersdorf AG, Research & Development, Unnastrasse 48, 20245 Hamburg, Germany

<sup>5</sup>Co-first author

\*Correspondence: jan.siemens@pharma.uni-heidelberg.de

<http://dx.doi.org/10.1016/j.cell.2015.01.022>

## SUMMARY

Sensitization of the capsaicin receptor TRPV1 is central to the initiation of pathological forms of pain, and multiple signaling cascades are known to enhance TRPV1 activity under inflammatory conditions. How might detrimental escalation of TRPV1 activity be counteracted? Using a genetic-proteomic approach, we identify the GABA<sub>B1</sub> receptor subunit as bona fide inhibitor of TRPV1 sensitization in the context of diverse inflammatory settings. We find that the endogenous GABA<sub>B</sub> agonist, GABA, is released from nociceptive nerve terminals, suggesting an autocrine feedback mechanism limiting TRPV1 sensitization. The effect of GABA<sub>B</sub> on TRPV1 is independent of canonical G protein signaling and rather relies on close juxtaposition of the GABA<sub>B1</sub> receptor subunit and TRPV1. Activating the GABA<sub>B1</sub> receptor subunit does not attenuate normal functioning of the capsaicin receptor but exclusively reverts its sensitized state. Thus, harnessing this mechanism for anti-pain therapy may prevent adverse effects associated with currently available TRPV1 blockers.

## INTRODUCTION

Pathological forms of pain are usually triggered by injury or inflammation of peripheral sensory neurons of the pain pathway. A diverse set of inflammatory stimuli can sensitize nociceptive neurons to promote pain hypersensitivity. As a consequence, inhibition of individual inflammatory pathways as a means to attenuate pain is a problematic approach for drug development, as parallel signaling cascades are still able to drive pathological, pain-promoting sensitization (Gold and Gebhart, 2010).

One receptor that has been found to serve as downstream integrator of many inflammatory pathways and thus holds great hope for anti-pain therapy is the capsaicin receptor TRPV1 (Caterina et al., 1997; Tominaga et al., 1998).

Next to its physiological function as a detector of noxious stimuli, a large body of literature attests to a crucial pathological

role for TRPV1. Importantly, inflammatory sensitization leads to dramatically reduced activation thresholds of TRPV1, producing hyperalgesia and pain hypersensitivity. Indeed, *Trpv1*<sup>-/-</sup> animals completely lack thermal hyperalgesia, confirming TRPV1's central role as integrator of disparate inflammatory pathways (Caterina et al., 2000; Davis et al., 2000).

TRPV1's central role in pathological forms of pain has sparked intense efforts to develop TRPV1 antagonists. However, blocking TRPV1 activity per se leads to impaired noxious heat sensation and produces hypothermia, attesting to a role of this temperature-gated receptor in homeostatic thermoregulation. Therefore, preventing and counteracting sensitization of the receptor while leaving basal TRPV1 activity untouched has been suggested to be a more promising avenue for pain therapy (Szalai and Sheta, 2012; Vay et al., 2012; Woolf, 2010).

Combining mouse genetics with proteomics, here we report the identification of the GABA<sub>B1</sub> receptor as a modulator of TRPV1 sensitization.

Canonical GABA<sub>B</sub> signaling requires GABA<sub>B1</sub> and GABA<sub>B2</sub> receptor subunits to orchestrate activation of inhibitory G proteins (Gassmann et al., 2004; Jones et al., 1998; Kaupmann et al., 1998; Kuner et al., 1999; White et al., 1998). Surprisingly, the GABA<sub>B2</sub> subunit is dispensable for mediating inhibition of TRPV1 sensitization, demonstrating that the underlying mechanism diverges from canonical GABA<sub>B</sub> receptor signaling.

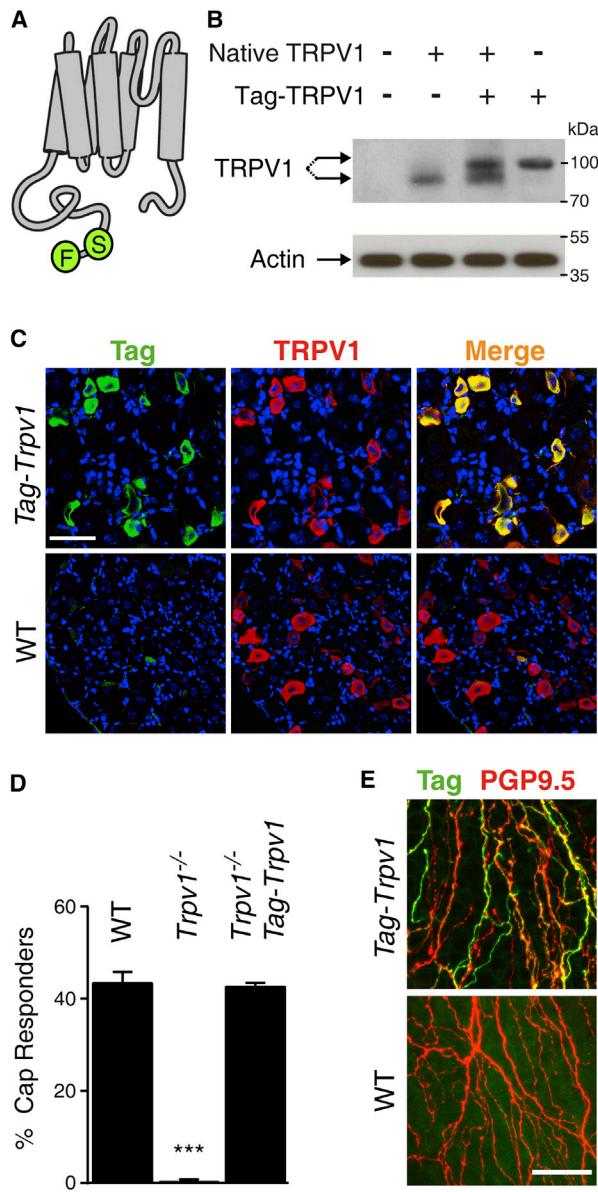
Collectively, our results establish an endogenous GABA/GABA<sub>B1</sub> feedback mechanism that keeps TRPV1-mediated pain hypersensitivity in check.

## RESULTS

### Transgenic Tag-TRPV1 Recapitulates Native TRPV1 Function

Given the complex integrative function of the capsaicin receptor, we reasoned that analysis of TRPV1 accessory proteins would shed light on unrecognized modulation of the receptor. Therefore, we utilized a genetic-biochemical approach to identify components of TRPV1 protein complexes isolated from native sensory tissue of transgenic animals.

To achieve efficient and specific biochemical purification of TRPV1 protein complexes for downstream analysis by mass spectrometry, we generated mice expressing an affinity-tagged



**Figure 1.** Tag-TRPV1 Reproduces Native TRPV1 Expression and Function

- (A) Cartoon depicting Tag-TRPV1.
- (B) Western blot of solubilized membrane fractions of DRG derived from WT, *Trpv1*<sup>-/-</sup>, and *Trpv1*<sup>-/-</sup>;Tag-*Trpv1* mice.
- (C) Immunostaining of DRG derived from WT and Tag-*Trpv1* mice with anti-Flag (green) and anti-TRPV1 (red) antibodies reveals that Tag-TRPV1 recapitulates the expression profile of the native receptor.
- (D) Tag-TRPV1 expression rescues capsaicin responses in *Trpv1*<sup>-/-</sup> mice. Cultured DRG neurons obtained from WT, *Trpv1*<sup>-/-</sup>, and *Trpv1*<sup>-/-</sup>;Tag-*Trpv1* mice were challenged with capsaicin and the proportion of responders assessed by calcium imaging. Error bars represent SEM.
- (E) Immunostaining of Tag-TRPV1 (anti-Flag antibody, green) and the pan-neuronal marker PGP9.5 (red) of corneas derived from WT and Tag-*Trpv1* mice reveals that Tag-TRPV1 protein is transported to peripheral sensory endings. Scale bars, 50 μm.

See also Figure S1.

version of TRPV1 (Tag-TRPV1, Figure 1A). Western blotting of protein extracts derived from sensory ganglia (dorsal root ganglia, DRG) of transgenic animals revealed the expected size of the tagged receptor protein, which is slightly shifted toward higher molecular weight compared to native (untagged) TRPV1 (Figure 1B).

To assure identification of physiological relevant TRPV1 protein complex components, we first verified cell-type-specific expression and functionality of the tagged receptor in BAC transgenic animals. Tag-TRPV1 expression recapitulated the native profile in DRG (Figures 1C, S1A, and S1B). Accordingly, the tagged receptor rescued capsaicin responses in DRG sensory neurons derived from transgenic animals crossed onto the TRPV1 knockout background (Figures 1D and S1C). Importantly, the Tag did not impair trafficking of TRPV1 to peripheral and central nerve terminals (Figures 1E and S1D).

Finally, we tested TRPV1-dependent nociceptive behavior. Whereas we observed normal responses to heat stimuli in a transgenic mouse line expressing close to physiological levels of Tag-TRPV1, a second line, expressing high levels of the tagged receptor (Figure S1E), exhibited significantly reduced pain thresholds (Figure S1F). These findings not only prove the functionality of the tagged receptor *in vivo* but also underscore TRPV1's predominance in mediating heat hyperalgesia.

### Mass Spectrometry Reveals the GABA<sub>B1</sub> Subunit as a Constituent of TRPV1 Protein Complexes

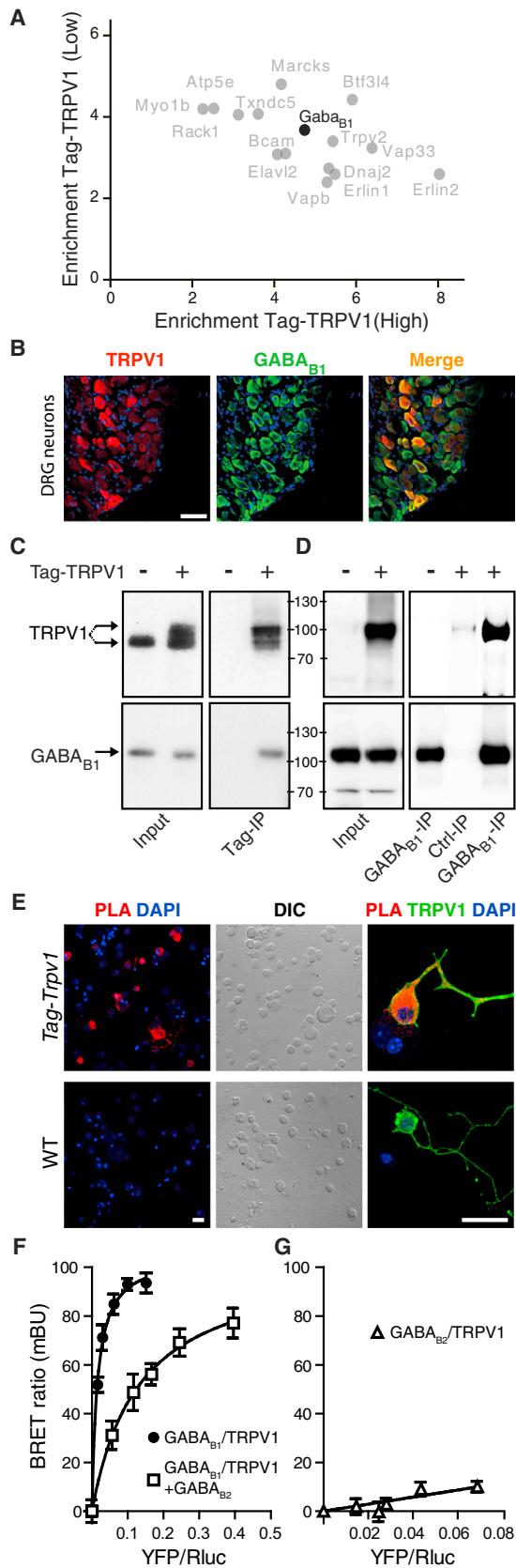
Next, utilizing the Strep or Flag moiety of the tagged receptor, we affinity purified protein complexes extracted from DRG of transgenic animals expressing both the tagged and the native form of the receptor. Wild-type animals devoid of the tagged receptor served as control.

Mass spectrometry revealed the presence of TRPV2—a TRP family member previously shown to interact with TRPV1 (Rutter et al., 2005)—in purified protein complexes, further validating our approach. The GABA<sub>B1</sub> receptor subunit was identified as one of the most abundant components of native TRPV1 protein complexes, regardless of Tag-TRPV1 expression level (Figures 2A and S2A and Table S1).

Interestingly, GABA<sub>B</sub> receptors in the CNS have long been known to mediate analgesia (Sokal and Chapman, 2003). However, GABA<sub>B</sub> mechanisms specifically targeting the peripheral nociceptive pathway have not been identified. Surprisingly, GABA<sub>B</sub> receptors are robustly expressed in peripheral nociceptive neurons (Charles et al., 2001; Towers et al., 2000), suggesting an unrecognized role for GABA<sub>B</sub> in the pain pathway. We were able to confirm these results and found that most, if not all, TRPV1-positive sensory neurons express the GABA<sub>B1</sub> subunit (Figure 2B).

We validated the presence of GABA<sub>B1</sub> in TRPV1 protein complexes by immunoprecipitation (IP) of Tag-TRPV1 or GABA<sub>B1</sub> (Figures 2C and 2D).

To further confirm the close proximity of the two transmembrane receptors in their native environment, we carried out proximity ligation assays (PLA). In this assay, a fluorescent signal is generated when two proteins of interest either physically interact directly or coexist within close molecular distance (Söderberg et al., 2006). We therefore cultured DRG neurons of Tag-*Trpv1*



### Figure 2. GABA<sub>B1</sub> and TRPV1 Reside in Common Protein Complexes

(A) Mass spectrometric analysis and quantification of Flag-affinity-purified Tag-TRPV1 protein complexes derived from DRG of mice expressing either high or low levels of Tag-TRPV1. Dot plot shows significantly enriched proteins as log<sub>2</sub> fold ratio (Tag-TRPV1/WT). The analysis reveals specific and robust enrichment for GABA<sub>B1a</sub> protein.

(B) Immunostaining of DRG of WT mice with anti-TRPV1 (red) and anti-GABA<sub>B1</sub> (green) antibodies shows co-expression of both receptors in nociceptive neurons. Scale bar, 50  $\mu$ m.

(C and D) Immunoprecipitations of solubilized protein extracts derived from DRG of *Tag-Trpv1* or WT control animals using immobilized anti-Flag antisera (C), anti-GABA<sub>B1</sub>, or control antisera (D). Western blots are probed with anti-GABA<sub>B1</sub> (bottom) or anti-TRPV1 (top panels in C) or anti-Flag antisera (top panels in D). GABA<sub>B1</sub> only co-elutes from affinity resin in the presence, but not in the absence, of TRPV1 (bottom right panel in C). Note that DRG neurons predominantly express the GABA<sub>B1a</sub> isoform and little or no GABA<sub>B1b</sub>. Similarly, TRPV1 is specifically enriched in isolated GABA<sub>B1</sub> protein complexes (top right panel in D). Note that endogenous (untagged) TRPV1 protein is precipitated together with Tag-TRPV1, confirming the biochemical isolation of native tetrameric TRPV1 protein complexes (top right panel in C).

(E) PLA of dissociated DRG neurons of *Tag-Trpv1* and WT animals. Antibodies against Flag and GABA<sub>B1</sub> were used to detect proximity of the two receptors. PLA signal (red) is only present in cell bodies and neurites derived from *Tag-Trpv1* animals. Scale bar, 25  $\mu$ m.

(F) BRET assay using Rluc-TRPV1 and GABA<sub>B1</sub>-YFP fusion proteins in the presence or absence of GABA<sub>B2</sub>, showing that TRPV1 and GABA<sub>B1</sub> strongly interact and that GABA<sub>B2</sub> competes with TRPV1 for GABA<sub>B1</sub> binding. Error bars represent SEM.

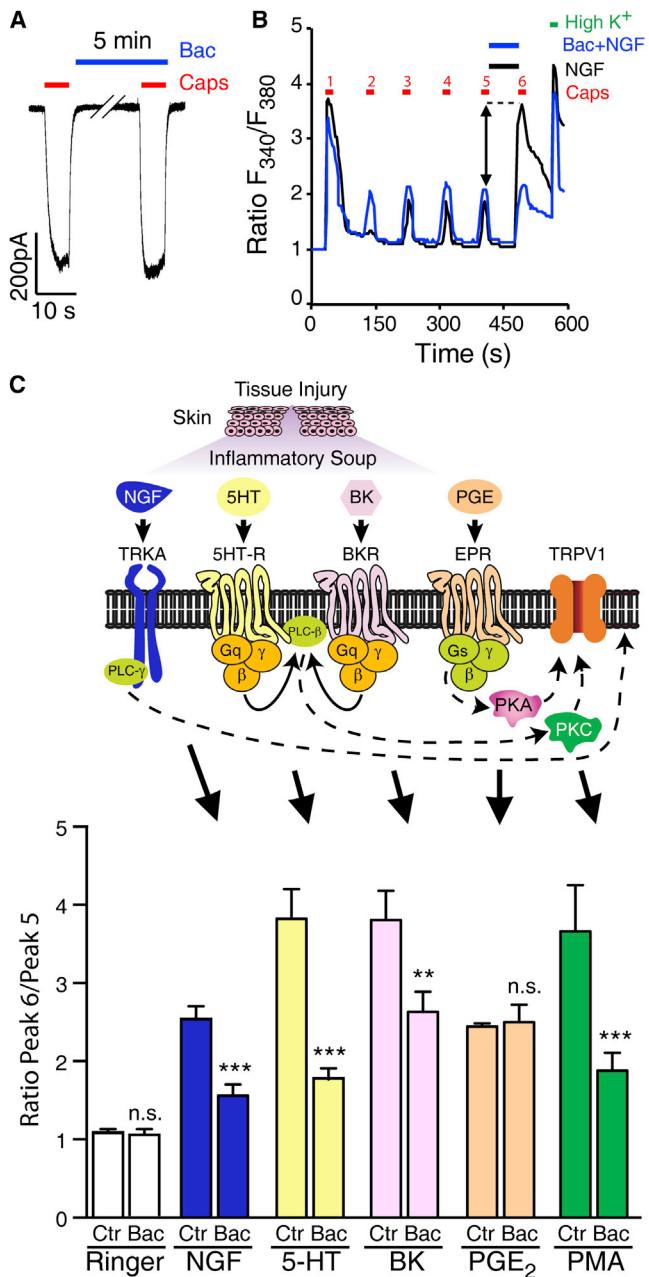
(G) No interaction was detected between Rluc-TRPV1 and GABA<sub>B2</sub>-YFP. Error bars represent SEM.  
See also Figure S2 and Table S1.

mice and wild-type control animals and performed PLA using anti-Flag and anti-GABA<sub>B1</sub> antibodies. A robust PLA signal was detected in *Tag-Trpv1* neurons, but not in WT controls, demonstrating that a TRPV1-GABA<sub>B1</sub> complex is present in intact sensory neurons and their projecting neurites (Figure 2E).

Intriguingly, we were not able to detect the GABA<sub>B2</sub> subunit either by mass spectrometry or by western blotting of TRPV1 protein complexes (Figures S2B and S2C). Similarly, KCTD proteins—found to constitutively interact with the GABA<sub>B2</sub> receptor subunit (Schwenk et al., 2010)—were absent, and only GABA<sub>B1</sub> selectively co-purified with TRPV1 complexes.

Additionally, when co-expressing TRPV1, GABA<sub>B1</sub>, and GABA<sub>B2</sub>, lower amounts of GABA<sub>B1</sub> protein were detected in TRPV1 IPs when compared to samples containing TRPV1 and GABA<sub>B1</sub> alone (Figures S2D and S2E). These findings suggested that TRPV1 and GABA<sub>B2</sub> might compete for binding to GABA<sub>B1</sub>. To further examine this possibility, we carried out bioluminescence resonance energy transfer (BRET) assays (Ayoub and Pfleger, 2010) and expressed *Renilla Luciferase* BRET donor (Rluc-TRPV1) and YFP BRET acceptor (GABA<sub>B1</sub>-YFP, GABA<sub>B2</sub>-YFP, or YFP-TRPV1) fusion proteins in a heterologous cell expression system. While GABA<sub>B</sub> BRET probes have been characterized previously (Adelfinger et al., 2014), we first confirmed the functionality of the TRPV1 BRET probes (Figures S2F and S2G).

The BRET assay confirmed close association of TRPV1 and GABA<sub>B1</sub> (Figure 2F). Again, no interaction was detectable between TRPV1 and GABA<sub>B2</sub> (Figure 2G). Moreover, similar to the co-IP experiments (Figures S2D and S2E), co-expression



**Figure 3. Effect of Baclofen on Native TRPV1 Receptors in Cultured Sensory Neurons**

(A) Electrophysiological recordings of cultured DRG neurons. TRPV1 currents elicited by capsaicin (500 nM) were not affected by application of 100  $\mu$ M baclofen. Rise time (10%–90%, in s): control, 6.5  $\pm$  1.8; Baclofen, 5.8  $\pm$  1.3. Decay ( $\tau$ , s): control, 7.6  $\pm$  1.5; Baclofen, 6.8  $\pm$  1.8.

(B)  $\text{Ca}^{2+}$  responses (measured as normalized fluorescence ratios) of DRG neurons stimulated repetitively with capsaicin (100 nM). NGF (100 ng/ml) sensitized TRPV1 activity resulting in a larger response magnitude. Co-application of baclofen (100  $\mu$ M) with NGF resulted in inhibition of TRPV1 sensitization.

(C) Top: simplified cartoon showing the release of inflammatory mediators during tissue injury. Components of the “inflammatory soup” activate diverse receptor types, which signal via PKC, PKA, and lipid modifiers, such as PLC, to mediate TRPV1 sensitization. Bottom: quantification of averaged fluorescence

of GABA<sub>B2</sub> interfered with the interaction of Rluc-TRPV1 and GABA<sub>B1</sub>-YFP (Figure 2F), shown by the increased amount of BRET acceptor (GABA<sub>B1</sub>-YFP) required to reach 50% of the maximal BRET signal (Figures 2F, S2H, and S2I).

In summary, GABA<sub>B1</sub> and TRPV1 form a protein complex that appears to lack the G protein signaling subunit GABA<sub>B2</sub>.

### GABA<sub>B</sub> Signaling Counteracts Sensitization of TRPV1

Given GABA<sub>B</sub>’s established role in modulating KIR3 and Cav channels (Padgett and Slesinger, 2010), it is conceivable that GABA<sub>B</sub>’s known analgesic property could stem in part from inhibiting TRPV1 activity.

Different to its effect on KIR3 channels (Figure S3A), neither calcium imaging nor electrophysiological recordings revealed any influence of the GABA<sub>B</sub> agonist baclofen on TRPV1 currents elicited by capsaicin (Figures 3A and S3B). Tachyphylaxis, a form of rapid TRPV1 desensitization, was also not affected by baclofen application (Figure S3C).

Under inflammatory conditions, multiple pathways sensitize TRPV1. Paramount among the different inflammatory sensitizers is NGF, activating parallel TRKA receptor signaling cascades that converge on the capsaicin receptor to enhance its sensitivity.

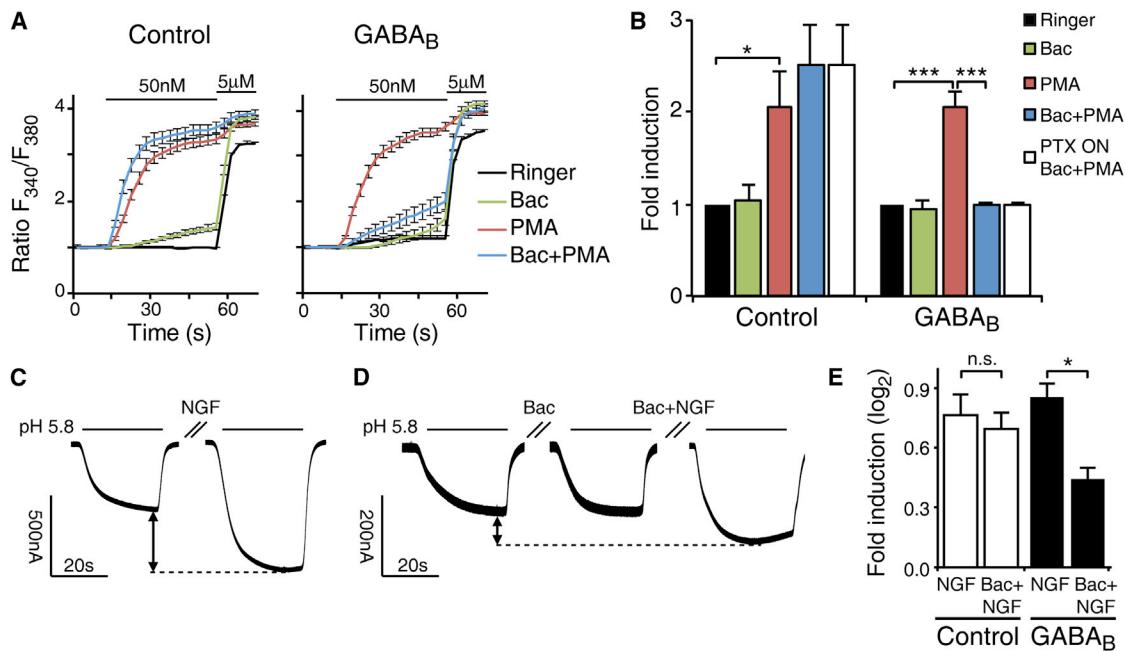
We thus tested whether NGF-mediated TRPV1 sensitization is modulated by GABA<sub>B</sub> activation. To this end, we monitored nociceptive neuron populations for NGF-enhanced TRPV1 activity using calcium imaging (Bonnington and McNaughton, 2003). Indeed, pre-incubation of sensory neurons with baclofen (but not saline) robustly blocked TRPV1 sensitization (Figures 3B, 3C, and S3D).

Apart from the receptor tyrosine kinase TRKA, several G-protein-coupled receptors (GPCRs) mediate TRPV1 sensitization (Figure 3C). Notorious in this regard are bradykinin- and serotonin-triggered GPCR cascades (Chuang et al., 2001; Huang et al., 2006). Strikingly, baclofen was equally effective in inhibiting TRPV1 sensitization induced by both inflammatory mediators (Figures 3C, S3E, and S3F).

Given that the initial signaling events controlled by the two GPCRs are quite different to that of TRKA, these results suggest that GABA<sub>B</sub> exerts its effect at a converging point downstream of the different pathways, potentially at TRPV1 itself. In agreement with this hypothesis, bypassing upstream inflammatory receptor signaling by direct pharmacological PKC activation also resulted in baclofen-reversible TRPV1 sensitization (Figure 3C).

The inflammatory prostaglandin PGE<sub>2</sub> mediates TRPV1 sensitization largely via a G<sub>s</sub>-coupled/PKA pathway (Gu et al., 2003; Lopshire and Nicol, 1998; Moriyama et al., 2005). Intriguingly, baclofen did not inhibit PGE<sub>2</sub>-mediated TRPV1 sensitization (Figure 3C). Canonical GABA<sub>B</sub>-G<sub>i/o</sub> coupling would be expected to effectively counteract G<sub>s</sub>-mediated sensitization. In agreement

ratios as shown in (B). Bar graphs represent the ratios between peaks 6 and 5 (after and before application of inflammatory agents) in the presence or absence of baclofen (100  $\mu$ M). Baclofen effectively attenuated TRPV1 sensitization induced by NGF (100 ng/ml), serotonin (100  $\mu$ M), bradykinin (10 nM), and PMA (1  $\mu$ M), but not by PGE<sub>2</sub> (1  $\mu$ M). Cells treated with vehicle or baclofen served as a negative control. Error bars represent SEM. See also Figure S3.



**Figure 4. Reversal of TRPV1 Sensitization by GABA<sub>B</sub> Is Recapitulated in Cellular Expression Systems In Vitro**

(A) Calcium responses elicited by 50 nM capsaicin pulses were assessed in TRPV1-HEK293 cells with or without GABA<sub>B</sub> after incubation with either baclofen (100  $\mu$ M), PMA (1  $\mu$ M), or the combination of PMA+baclofen as indicated. Depicted are representative ratiometric traces from one experiment ( $n = 50$  cells). PMA promotes PKC-induced TRPV1 sensitization, and baclofen inhibits sensitization only in the presence of GABA<sub>B</sub>. Error bars represent SEM.

(B) Quantification of experiments shown in (A). PTX (500 ng/ml) did not abrogate the baclofen effect, indicating that G<sub>i/o</sub> proteins are not involved in counteracting TRPV1 sensitization by GABA<sub>B</sub>.

(C and D) Electrophysiology recordings of oocytes injected with *Trpv1*, *TrkA*, and *Gaba<sub>B</sub>*, following TRPV1 activation with pH 5.8, before and after treatment with NGF (100 ng/ml) (C) and NGF + baclofen (D).

(E) Quantification of the experiments shown in (C) and (D); sensitization was measured as log<sub>2</sub> ratio of capsaicin-elicited steady-state currents before and after drug treatment. TRPV1 sensitization mediated by NGF is attenuated by baclofen only in the presence of GABA<sub>B</sub>. Error bars represent SEM.

See also Figure S4.

with the apparent absence of the GABA<sub>B2</sub> subunit from TRPV1-GABA<sub>B1</sub> complexes, this result further pointed at a potentially unrecognized GABA<sub>B</sub> mechanism modulating TRPV1 sensitization independent of G<sub>i/o</sub> signaling.

#### Non-Canonical GABA<sub>B</sub> Signaling Counteracts TRPV1 Sensitization

To test directly whether G<sub>i/o</sub> protein is involved in mediating the observed GABA<sub>B</sub> effect, we pre-incubated sensory neurons with pertussis toxin (PTX). This selective and potent blocker of G<sub>i/o</sub> proteins did not attenuate baclofen's effect on TRPV1 sensitization (Figures S4A and S4B), whereas another G<sub>i/o</sub>-signaling cascade present in DRG neurons was effectively inhibited by PTX (Figures S4C and S4D), thereby serving as a positive control.

The positive allosteric modulator CPG7930 potentiates GABA<sub>B</sub> receptor signaling by binding to the GABA<sub>B2</sub> subunit (Binet et al., 2004). However, inhibition of TRPV1 sensitization was not significantly enhanced by CPG7930 (Figure S4E) at concentrations that robustly potentiated GABA<sub>B</sub>-induced KIR3 currents (Figure S4F), again pointing toward a non-canonical GABA<sub>B</sub> pathway.

Next, we tested GABA<sub>B</sub>'s modulatory effect on TRPV1 sensitization in two heterologous expression systems.

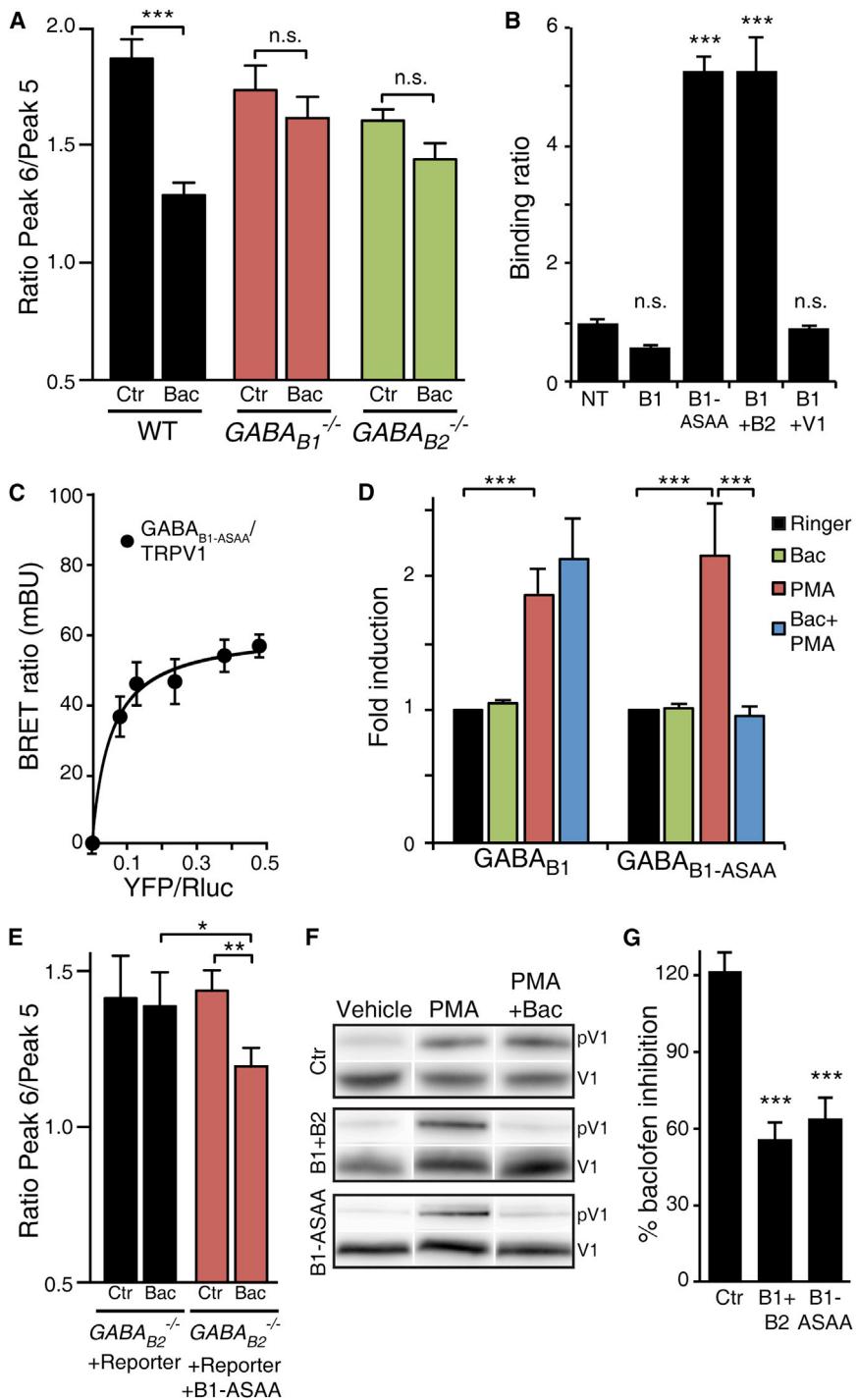
We first recapitulated PKC-induced TRPV1 sensitization in a TRPV1-HEK293 cell line (Siemens et al., 2006). Pre-incubating the cells with baclofen in the absence of GABA<sub>B</sub> had no effect and did not abrogate PKC-mediated TRPV1 sensitization. However, co-expressing GABA<sub>B</sub> completely reverted TRPV1 hyperactivity upon baclofen incubation (Figures 4A and 4B). Similar to our results in native sensory neurons, PTX did not prevent baclofen to exert inhibition of TRPV1 sensitization (Figure 4B).

Additionally, we reconstituted the NGF sensitization pathway and expressed the TRKA receptor together with TRPV1 and GABA<sub>B</sub> in *Xenopus* oocytes. Again, we found that baclofen was able to significantly reduce NGF sensitization in a GABA<sub>B</sub>-dependent manner (Figures 4C–4E).

In agreement with the observed close molecular proximity of the two receptors, these results demonstrate a direct inhibitory effect of GABA<sub>B</sub> receptors on TRPV1 sensitization, independent of classical G<sub>i/o</sub> signaling downstream of GABA<sub>B2</sub>.

#### GABA<sub>B1</sub> Reverts TRPV1 Sensitization Independent of GABA<sub>B2</sub> Signaling

In order to examine the role of the two GABA<sub>B</sub> receptor subunits in more detail, we next employed GABA<sub>B</sub> receptor knockout mice. A conditional "floxed" GABA<sub>B1</sub> mouse line (*Gaba<sub>B1</sub>*<sup>fl/fl</sup>; Haller et al., 2004) was crossed to *Trpv1*-Cre mice (Cavanaugh et al.,



**Figure 5. Surface Expression of  $GABA_{B1}$  in the Absence of  $GABA_{B2}$  Is Sufficient to Mediate Inhibition of TRPV1 Sensitization**

(A) Quantification of capsaicin responses in DRG neurons from *Trpv1-Cre;Gaba<sub>B1</sub><sup>fl/fl</sup>* ( $Gaba_{B1}^{-/-}$ ),  $Gaba_{B2}^{-/-}$ , and WT control mice in the presence or absence of baclofen (100  $\mu$ M) following incubation with serotonin (100  $\mu$ M). Error bars represent SEM.

(B) Cell surface expression of  $GABA_{B1}$  and  $GABA_{B1-ASAA}$  in the presence or absence of  $GABA_{B2}$  or TRPV1 was assessed by measuring the binding of [<sup>3</sup>H]CPGP54626 in intact CHO cells. Unspecific binding to non-transfected (NT) cells is set to 1. Error bars represent SEM.

(C) BRET experiment demonstrating that RLUC-TRPV1 and surface-localized  $GABA_{B1-ASAA}$ -YFP fusion proteins interact. Error bars represent SEM.

(D) Calcium responses in TRPV1-HEK293 cells transfected with  $GABA_{B1}$  or  $GABA_{B1-ASAA}$  after incubation with either baclofen (100  $\mu$ M), PMA (1  $\mu$ M), or the combination of PMA+baclofen. Error bars represent SEM.

(E) Quantification of serotonin-sensitized capsaicin responses of cultured  $Gaba_{B2}^{-/-}$  DRG neurons transfected with a fluorescent reporter (tomato) in the presence or absence of  $GABA_{B1-ASAA}$  plasmid, incubated with baclofen (100  $\mu$ M). Error bars represent SEM.

(F) Representative western blot showing phosphorylated (pV1) and total TRPV1 protein (V1) of HEK cells transfected with control plasmid,  $GABA_{B1+B2}$ , or  $GABA_{B1-ASAA}$  alone after treatment with PMA, PMA+baclofen, or vehicle. Three biological replicates are loaded onto the gel, but only one representative lane for each condition is shown.

(G) Quantification of (F). Phosphorylated TRPV1 intensity values were normalized to total TRPV1 intensities and expressed as percent of baclofen-mediated inhibition (PMA+baclofen/PMA). Error bars represent SEM.

See also Figure S5.

ure 5A), demonstrating that the  $GABA_{B2}$  subunit is required for mediating TRPV1 modulation, despite its apparent absence from TRPV1 protein complexes and the dispensability of  $G_{i/o}$  signalling downstream of the  $GABA_{B2}$  subunit.

It is well established that  $GABA_{B1}$  does not translocate to the cell surface autonomously, but the  $GABA_{B2}$  subunit mediates trafficking of the heteromeric receptor complex by masking an ER

retention signal present in the  $GABA_{B1}$  protein (Margera-Mitrovic et al., 2000; Pagano et al., 2001). We therefore asked whether TRPV1 promotes cell surface translocation of  $GABA_{B1}$  independent of the  $GABA_{B2}$  subunit. Analyzing cell surface binding of a  $GABA_{B1}$ -specific radioligand [<sup>3</sup>H]CPGP54626 (Galvez et al., 2001), we find that TRPV1, different to  $GABA_{B2}$ , does not promote cell surface translocation of  $GABA_{B1}$  (Figure 5B).

2011) to specifically ablate the  $GABA_{B1}$  subunit in TRPV1-positive sensory neurons (Figure S5A). As expected, inhibition of TRPV1 sensitization by baclofen was absent in sensory neurons derived from conditional *Trpv1-Cre;Gaba<sub>B1</sub><sup>fl/fl</sup>* mice, attesting to baclofen's high specificity and selectivity for the  $GABA_{B1}$  subunit (Figure 5A). In DRG neurons derived from  $Gaba_{B2}^{-/-}$  mice (Gassmann et al., 2004), the baclofen effect was also abrogated (Fig-

Mutating the ER-retention signal encoded by the peptide sequence “RSRR” to “ASAA” allows surface translocation of GABA<sub>B1</sub> independent of GABA<sub>B2</sub> without changing GABA<sub>B1</sub>’s signaling capability (Margeta-Mitrovic et al., 2000; Pagano et al., 2001). Interestingly, this constitutively surface-localized GABA<sub>B1</sub> variant, GABA<sub>B1-ASAA</sub> (Figure 5B), interacts with TRPV1 similar to its wild-type counterpart (Figure 5C), suggesting that the GABA<sub>B1</sub>-TRPV1 protein complex can form at the cell surface.

Expectedly, ER-retained native GABA<sub>B1</sub> protein was not in the position to mediate inhibition of TRPV1 sensitization when expressed in the absence of the GABA<sub>B2</sub> subunit (Figure 5D). Strikingly, surface-localized GABA<sub>B1-ASAA</sub> alone robustly inhibited TRPV1 sensitization in HEK293 cells (Figure 5D). Moreover, GABA<sub>B1-ASAA</sub> reinstalled baclofen-mediated inhibition of TRPV1 sensitization in GABA<sub>B2</sub>-deficient sensory neurons (Figures 5E, S5B, and S5C), demonstrating the ability of GABA<sub>B1</sub> to signal to TRPV1 directly and independently of the GABA<sub>B2</sub> subunit.

PKC phosphorylation of TRPV1 sensitizes the receptor (Mandadi et al., 2006) and is a convergence point downstream of NGF, bradykinin, and serotonin signaling. We therefore examined whether PMA-induced PKC phosphorylation of TRPV1 is attenuated by GABA<sub>B1</sub> signaling. Indeed, we find that GABA<sub>B1-ASAA</sub> alone (similar to the native GABA<sub>B1+B2</sub> receptor pair) effectively inhibits PKC phosphorylation of TRPV1 (Figures 5F and 5G).

### GABA<sub>B</sub> Reverts TRPV1 Hypersensitivity in Models of Inflammatory and Neuropathic Pain

Given the robust cellular effects of GABA<sub>B1</sub> on TRPV1 sensitization, we next tested whether these results translate into GABA<sub>B1</sub>-mediated inhibition of pain hypersensitivity in animal models. One nociceptive assay that allows robust measurement of TRPV1-dependent heat hyperalgesia is the Hargreaves assay. In this assay, inflammatory stimuli, such as NGF and bradykinin, decrease the paw withdrawal latency in a TRPV1-dependent manner (Chuang et al., 2001). Indeed, we observed significant inhibition of sensitization in baclofen-injected mice compared to mice that solely received NGF. No effects were detected in *Trpv1*<sup>-/-</sup> animals (Figure S6A). To rule out any potential central effects of baclofen (Sokal and Chapman, 2003), we slightly modified the testing paradigm and induced inflammatory hyperalgesia in both hind paws of the same animal (Figure S6B). We found robust and dose-dependent inhibition of sensitization only in the ipsilateral paw that had received baclofen, but not in the contralateral paw (Figures 6A and S6C). Moreover, baclofen’s inhibitory effect was completely blunted in conditional *Trpv1-Cre;Gaba<sub>B1</sub>*<sup>ff</sup> mice (Figure 6B). These findings are in agreement with our cellular data (Figure 5A) and demonstrate that GABA<sub>B1</sub>’s beneficial effect on heat hyperalgesia is mediated in peripheral TRPV1-positive nociceptors in a cell-autonomous fashion, rather than in higher-order pain-processing CNS areas.

Similarly, GABA<sub>B1</sub> receptor activation strongly attenuated serotonin- and bradykinin-triggered heat hyperalgesia but spared PGE<sub>2</sub>-triggered heat hyperalgesia (Figures 6A and S6D), an effect that was not further enhanced by the positive allosteric modulator CPG7930 that stimulates GABA<sub>B2</sub> (Figure S6E).

For therapeutic purposes, it would be beneficial to identify an intervention that is able to revert pre-existing pain hypersensitiv-

ity. Applying baclofen at a time point when NGF had already sensitized nociceptors reduced heat withdrawal thresholds (Figure S6F), establishing that GABA<sub>B1</sub> activation has the capacity for both prevention and reversal of capsaicin receptor sensitization.

Given the ability of the GABA<sub>B1</sub> subunit to inhibit sensitization triggered by multiple different inflammatory pathways, we asked whether it is also effective in attenuating hyperalgesia as a consequence of intraplantar CFA (complete Freund’s adjuvant) injection, a widely used model of persistent inflammatory pain. We found that intraplantar application of baclofen 2 days after inducing CFA-mediated inflammation significantly reduced heat hyperalgesia compared to vehicle-treated controls (Figure 6C). The beneficial effect of a single dose lasted for a minimum of 3 hr and had completely ceased 24 hr after application.

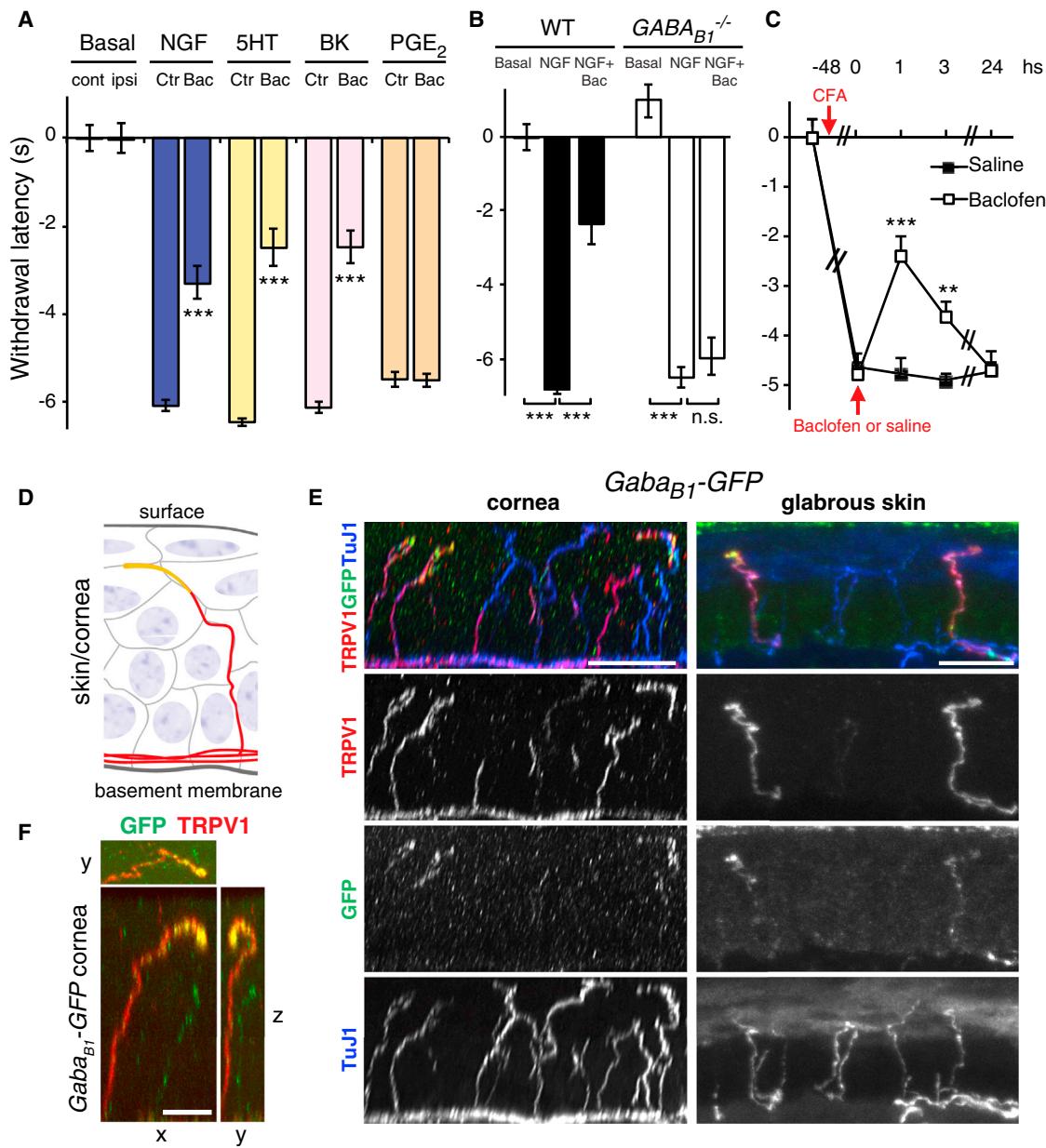
Similar to the cellular assays, our behavior experiments confirm that GABA<sub>B1</sub> activation specifically attenuates hypersensitivity of TRPV1 but did not impinge upon acute responsiveness of the receptor to either a heat stimulus, as tested in the Hargreaves assay (Figure S6G), or assessed by measuring acute capsaicin-induced nociceptive responses (Figure S6H).

### GABA<sub>B1</sub> and TRPV1 Co-Localize at Peripheral Nerve Endings

Our data suggest that GABA<sub>B1</sub> receptors mediate their anti-hyperalgesic effect in the periphery. Consequently, GABA<sub>B1</sub> would be expected to reside in free nerve endings in close proximity to TRPV1. To elucidate whether the two receptors indeed co-localize at nociceptor endings, we utilized transgenic *Gaba<sub>B1</sub>-GFP* reporter mice (Casanova et al., 2009). Similar to skin, the cornea is densely innervated by TRPV1-positive sensory fibers, attesting to strong reactions of the eye to inflammatory insults. We analyzed both skin and cornea tissue for co-expression of the two receptors in peripheral nerve endings that project from DRG and trigeminal (TG) sensory neurons, respectively. Not only did we find GABA<sub>B1</sub> to be present in peripheral nociceptive fibers, but we also observed substantial co-localization with TRPV1 in nerve terminals (Figures 6D–6F and S6I–S6K and Movie S1). In agreement with the close association of TRPV1 and GABA<sub>B1</sub> found in Tag-TRPV1-positive DRG neurons (Figure 2E), 89% of TRPV1-positive corneal fibers co-localize the two receptors at their terminals (Figure S6L).

### GABA Is Released from Peripheral Nociceptors and Modulates Pain Hypersensitivity

Interestingly, we found that local application of the competitive GABA<sub>B1</sub> receptor antagonist CPG52432 enhanced NGF-mediated TRPV1 sensitization and induced a small but significant increase in mouse pain behavior compared to NGF alone (Figure 7A). This result suggested that GABA, the native agonist of GABA<sub>A</sub> and GABA<sub>B</sub> receptors, is endogenously present at peripheral nerve endings to produce a basal GABA<sub>B</sub> receptor tone that regulates TRPV1 sensitivity. Indeed, in blister fluid obtained from human skin and extracted mouse corneal fluid, we find GABA at concentrations ( $161 \pm 42$  nM and  $335 \pm 54$  nM, respectively) sufficient to inhibit TRPV1 sensitization in cultured DRG neurons in a GABA<sub>B1</sub>-dependent manner (Figure 7B).



**Figure 6. GABA<sub>B1</sub> Activation in Peripheral Nociceptive Terminals Decreases Thermal Hyperalgesia**

(A) Thermal responses were measured using the Hargreaves test. NGF (2 µg/paw), serotonin (5HT, 100 nmol/paw), bradykinin (10 nmol/paw), or PGE<sub>2</sub> (1 nmol/paw) were injected into both mouse hind paws 30 min after unilateral administration of baclofen (Bac, 3 µg/paw) into the ipsilateral paw and vehicle (Ctr) into the contralateral paw. Bar graph shows the differences in mean paw withdrawal latency upon radiant heat stimulation relative to the basal (non-stimulated) withdrawal latency. Baclofen decreased thermal hyperalgesia for all inflammatory mediators, except PGE<sub>2</sub>. Basal withdrawal latencies were measured 24 hr before the experiment. Error bars represent SEM.

(B) Baclofen does not attenuate NGF-induced thermal hyperalgesia in conditional *Trpv1-Cre;GABA<sub>B1</sub><sup>fl/fl</sup>* mice (*GABA<sub>B1</sub><sup>-/-</sup>*). Error bars represent SEM.

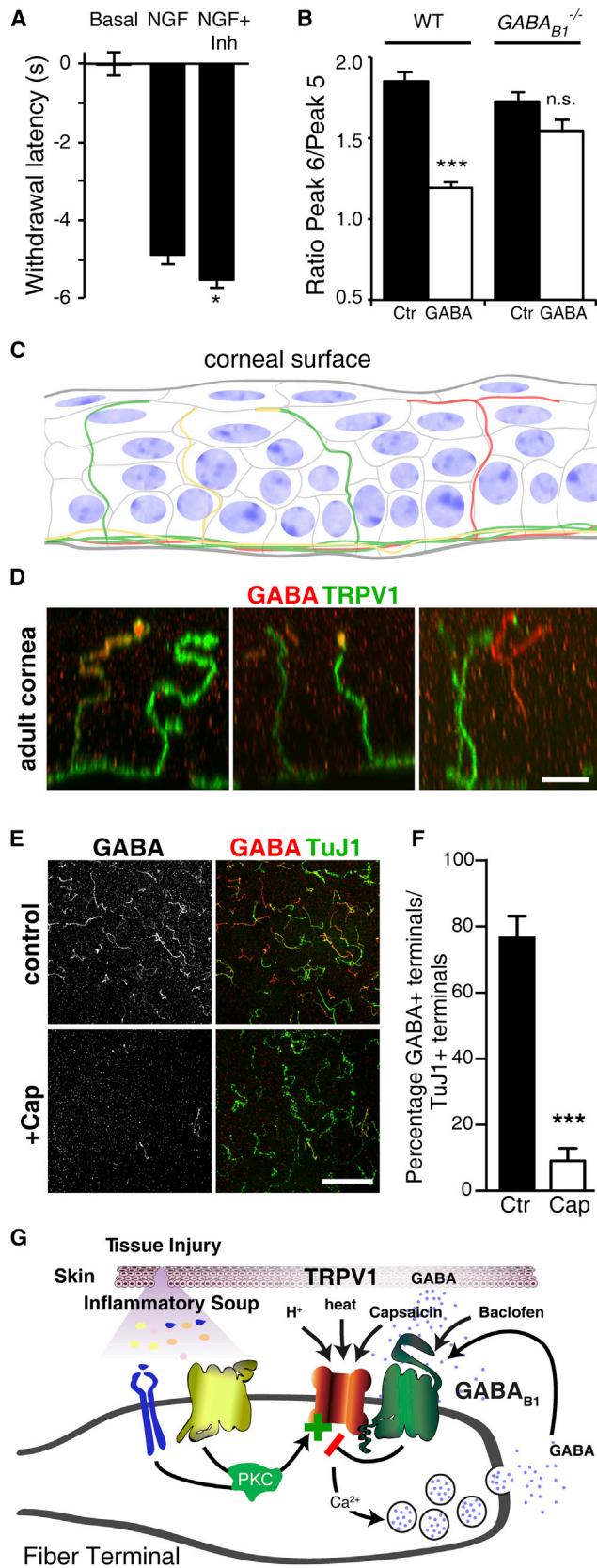
(C) Heat hyperalgesia 2 days after CFA induction is attenuated by baclofen (3 µg/paw) for a minimum of 3 hr when compared to mice that received vehicle only (saline). Error bars represent SEM.

(D) Cartoon depicting skin or cornea innervation shown in (E) and (F).

(E) Immunostaining of skin and cornea sections obtained from *GABA<sub>B1</sub>-GFP* transgenic mice and labeled with antisera for TRPV1 (red), GFP (green), and Tuj1 (blue). Scale bars, 20 µm.

(F) Higher magnification of a single corneal fiber. GABA<sub>B1</sub>-GFP localizes in TRPV1-positive fibers and is concentrated at the terminal close to the epithelial surface (demonstrated by orthogonal maximal projections in the x, y, and z dimensions). Scale bar, 10 µm.

See also Figure S6 and Movie S1.



**Figure 7. Peripheral Effects and Localization of GABA and Its Release from Peripheral Fibers**

(A) Withdrawal latencies in mice sensitized with NGF are significantly reduced by the application of the GABA<sub>B1</sub> inhibitor CPG52432 (17 nmol/paw) compared to mice that received vehicle. Error bars represent SEM.

(B) 200 nM GABA is sufficient to attenuate serotonin sensitization of TRPV1 in WT but not in *Trpv1-Cre;GABA<sub>B1</sub><sup>fl/fl</sup>* (*GABA<sub>B1</sub>*<sup>-/-</sup>) sensory neurons. Error bars represent SEM.

(C) Cartoon depicting cornea innervation shown in (D); nerve fibers emerge from the sub-basal nerve plexus and branch out toward the corneal surface. (D) Sensory fibers innervating the mouse cornea were immune labeled with antisera decorating GABA (red) and TRPV1 (green). GABA is either distributed throughout the length of the fibers (left) or concentrated at the terminals close to the corneal surface (middle) of TRPV1-positive fibers. Although most fibers contain GABA and TRPV1 conjointly, some fibers harbor GABA in the absence of TRPV1 (right). Scale bar, 10 μm.

(E) Corneas of WT mice either treated with capsaicin (10 μM) or vehicle (10 min, 32°C) and subsequently prepared for immunostainings using antisera recognizing GABA and TuJ1. GABA staining was strongly reduced in capsaicin-treated corneas compared to control, indicating that TRPV1 activation stimulates GABA release. Scale bar, 50 μm.

(F) Quantification of (E). Error bars represent SEM.

(G) Model illustrating inhibition of TRPV1 sensitization by GABA<sub>B1</sub> receptor signaling. TRPV1 is sensitized by a multitude of signaling cascades that are initiated in the context of inflammation. Several cascades converge on PKC, a kinase shown to robustly mediate TRPV1 sensitization via direct receptor phosphorylation. Sensitization augments TRPV1 channel activity, which leads to increased calcium influx and concomitant GABA release to attenuate TRPV1 sensitization via GABA<sub>B1</sub>. Activation of this autocrine loop dampens (or resolves) hyperactivity of TRPV1.

See also Figure S7.

Although GABA is well known for its presence in inhibitory CNS neurons, much less is known about its localization in the periphery. Using a GABA-specific antibody, we found GABA localized to the terminal endings of corneal nociceptors, many of which also express TRPV1 (Figures 7C, 7D, and S7A). Given that primary afferent nociceptors are glutamatergic, it was unexpected to find GABA at their peripheral endings. Presumably, in these fibers, GABA is stored in vesicles as many of the terminals also reacted with antibodies for vGAT, the vesicular GABA transporter (Figure S7B). Additionally, DRG and TG sensory neurons exhibit low Gad2 transcript levels (Figure S7C). Consistently, we find GABA to be present in DRG and TG neuronal cell bodies (Figure S7D), albeit the GABA labeling was less pronounced compared to that observed in TRPV1-positive sensory terminals (Figure 7D), likely reflecting efficient vesicular transport to distal terminals.

Given the juxtaposition of GABA vesicles and TRPV1 at peripheral nerve terminals, we tested whether TRPV1 activation and subsequent calcium influx would constitute an adequate stimulus to release GABA. Stimulating isolated corneas with capsaicin reduces GABA content at the terminals compared to corneas kept without capsaicin, strongly suggesting that GABA release has taken place (Figures 7E and 7F). The capsaicin-mediated release of GABA is specific to TRPV1 activation, as it was absent from capsaicin-stimulated corneas of *Trpv1*<sup>-/-</sup> mice (Figure S7E).

Taken together, these results support a model by which peripheral GABA limits (and resolves) TRPV1 sensitization via GABA<sub>B1</sub> receptors, directly at the site where painful stimuli are first encountered (Figure 7G).

## DISCUSSION

### Genetic-Proteomic Dissection of TRPV1 Modulation

Counteracting sensitization of TRPV1—without blocking acute channel activity—is a promising approach for developing analgesics. Using infusible peptides mimicking intracellular TRPV1 domains, it has been demonstrated that interference with TRPV1 sensitization has beneficial effects in models of inflammatory pain (Fischer et al., 2013). However, in contrast to a multitude of pathways sensitizing TRPV1, endogenous pathways inhibiting sensitization have remained elusive.

Combining a genetic-biochemical approach with quantitative mass spectrometry to probe the molecular environment of native TRPV1 receptors, we report here the identification of the GABA<sub>B1</sub> receptor as a modulator of TRPV1 sensitization.

### Non-Canonical GABA<sub>B</sub> Signaling Aborts TRPV1 Sensitization

In CNS neurons, most, if not all, GABA<sub>B</sub> receptor effects are mediated by coupling to inhibitory G<sub>i/o</sub>-type G proteins (Padgett and Slesinger, 2010). Activating G<sub>i/o</sub>-coupled signaling in sensory neurons is able to promote TRPV1 sensitization (Forster et al., 2009; Loo et al., 2012). Hence, it is difficult to reconcile the observed inhibitory GABA<sub>B</sub> effect on TRPV1 sensitization with a classical G<sub>i/o</sub>-coupled signaling cascade.

Multiple lines of evidence suggest that a GABA<sub>B1</sub> mechanism that is independent of GABA<sub>B2</sub> signaling targets TRPV1 hypersensitivity: (1) GABA<sub>B2</sub> appears to be absent from GABA<sub>B1</sub>-TRPV1 complexes; (2) neither PTX nor CPG7930— inhibiting G<sub>i/o</sub> activation or enhancing G protein coupling of GABA<sub>B2</sub>, respectively—had any effect on GABA<sub>B</sub>'s potential to attenuate TRPV1 sensitization; and (3) plasma-membrane-localized GABA<sub>B1</sub> can inhibit TRPV1 sensitization in the absence of GABA<sub>B2</sub>.

Our data suggest that a ternary complex encompassing TRPV1 and both GABA<sub>B</sub> subunits is not favored; rather, GABA<sub>B2</sub> competes with TRPV1 for GABA<sub>B1</sub> interaction. Interestingly, it has been shown that TRPV1 activation by capsaicin can promote dissociation of the two GABA<sub>B</sub> subunits (Laffray et al., 2007). Thus, it is tempting to speculate that TRPV1 activation promotes the formation of a surface-localized TRPV1-GABA<sub>B1</sub> liaison at the expense of a heteromeric GABA<sub>B1</sub>-GABA<sub>B2</sub> receptor complex, thereby priming the nociceptor to guard against sensitization.

### GABA<sub>B</sub> Inhibits Convergent PKC-Mediated TRPV1 Sensitization

How does the GABA<sub>B1</sub> receptor mediate its effect on TRPV1? Given the close proximity of the two proteins, it is possible that baclofen-induced conformational changes are directly transmitted onto the TRP ion channel to control and regulate its sensitization status. This would be reminiscent of dopamine receptors, which directly modulate GABA<sub>A</sub>-type ion channels, independent of a downstream signaling cascade (Liu et al., 2000).

Alternatively, a putative signaling effector downstream of GABA<sub>B1</sub> may inhibit TRPV1 sensitization. Signaling cascades downstream of GABA<sub>B1</sub> (and independent of GABA<sub>B2</sub>) have remained elusive. However, β-arrestin2 has been implicated to act downstream of GABA<sub>B</sub> receptors, independent of G<sub>i/o</sub> protein signaling (Lu et al., 2012). β-arrestin2 has also been

proposed to modulate tachyphylactic TRPV1 responses (Por et al., 2012), suggesting that this protein could, in principle, couple GABA<sub>B</sub> receptors to TRPV1. However, we did not detect β-arrestin2 in TRPV1-GABA<sub>B1</sub> protein complexes, nor did we find tachyphylactic TRPV1 responses to be modulated by GABA<sub>B</sub> receptor activation (Figure S3).

Different to the important functions for β-arrestins in other GPCR signaling cascades, their role in GABA<sub>B</sub>-mediated pathways is less clear (Perroy et al., 2003; Sudo et al., 2012), suggesting that β-arrestin2 is an unlikely candidate to mediate the modulatory action of GABA<sub>B1</sub> on TRPV1.

Different than the inflammatory pathways initiated by NGF, serotonin, or bradykinin, we find that PGE<sub>2</sub>-triggered sensitization of TRPV1 was not inhibited by the GABA<sub>B1</sub> receptor. The differential susceptibility to GABA<sub>B1</sub> receptor modulation coincides with different points of signal convergence on TRPV1; a common denominator of NGF, serotonin, and bradykinin signaling is the activation of PLC/PKC pathways that result in sensitization of the capsaicin receptor (Huang et al., 2006). By triggering PKC activation directly, we find that this branch of TRPV1 sensitization can be inhibited by GABA<sub>B1</sub> receptor activity, resulting in reduced TRPV1 phosphorylation.

In contrast, PGE<sub>2</sub> signaling leads to PKA-dependent phosphorylation of TRPV1 (Gu et al., 2003; Lopshire and Nicol, 1998; Moriyama et al., 2005), and its sensitizing effect was not attenuated by GABA<sub>B1</sub>. Phosphorylating TRPV1 at different sites may have different functional consequences despite a similar net effect on sensitization—recent studies demonstrate unique structural features of TRPV1 with an upper and lower gate that are independently engaged by different pro-algesic agents to open and modulate ion channel conductance (Cao et al., 2013). Thus, it is conceivable that differential phosphorylation by PKC and PKA mediates different types of TRPV1 gating, the first of which is susceptible to GABA<sub>B1</sub> receptor-mediated inhibition, whereas the latter is not.

Given that PKC is not the sole mediator of TRPV1 sensitization utilized by the inflammatory pathways susceptible to GABA<sub>B1</sub> inhibition, it is very well possible that other mechanisms are also targeted and that GABA<sub>B</sub> provides a broader protection (or “shielding”) against TRPV1 sensitization.

GABA<sub>B</sub> receptors trigger a surprising range of different cellular responses. Heteromeric receptor assembly, as well as the recent discovery of auxiliary GABA<sub>B</sub> subunits (Schwenk et al., 2010), has been proposed to explain contextual signaling diversity and challenge the classical view of GPCR organization and function. Our study further corroborates GABA<sub>B</sub> receptor signaling complexity and highlights a GABA<sub>B2</sub>-independent pathway that mitigates TRPV1-mediated pain hypersensitivity.

### Peripheral GABA Regulates Nociceptor Sensitization

It came as a surprise to find not only GABA<sub>B</sub> receptors but also their cognate agonist, GABA, localized at peripheral nerve terminals. Classically, vesicular GABA is a hallmark of inhibitory synapses of CNS neurons. Although it is known that the excitatory transmitter glutamate is present in peripheral nerves, where it may promote excitability and neurogenic inflammation (Miller et al., 2011), to our knowledge nothing similar has been reported for inhibitory transmitters.

Here, we provide evidence that, under physiological conditions, peripheral GABA limits TRPV1-mediated hyperalgesia. It will be interesting to ascertain whether the peripheral GABA<sub>B1</sub> feedback on TRPV1 is altered under pathological pain conditions. Exploiting this endogenous feedback system (e.g., by promoting peripheral GABA release or rendering a pharmacological GABA<sub>B1</sub> agonist non-permeable to the blood-brain barrier) may be a valuable route for anti-pain therapy, circumventing severe adverse effects associated with baclofen's dominant CNS activity.

Processing and modulation of painful signals have mainly been attributed to higher-order brain centers such as the dorsal spinal cord and beyond. Our model may—after all—not only inspire new approaches for developing TRPV1-centric pain therapeutics, but it also offers a fresh look at modulation of sensory input directly at the site of sensory transduction in nerve terminals, a paradigm that may also be relevant for other somatosensory modalities.

## EXPERIMENTAL PROCEDURES

### Generation of Tag-Trpv1 BAC Transgenic Mice

The SF (Strep-Flag) tag was seamlessly integrated into a mouse BAC clone (RP23-390G23) encoding the *Trvp1* genomic locus, and transgenic mice were obtained by pronuclear injection.

### Behavioral Studies

All animal experiments were in accordance with the local governing bodies. Thermal pain was assessed by measuring the response latency to a radiant heat stimulus focused onto the plantar surface of the paw (Hargreaves assay).

### Biochemical Protein Complex Purification and Mass Spectrometry Analysis

Following plasma membrane fractionation of DRG, protein complexes were affinity isolated using Anti-Flag magnetic beads. Liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis was performed with in-solution digested affinity purified protein samples on a Q Exactive mass spectrometer (Thermo Scientific). Label-free quantitation (LFQ) was performed using MaxQuant Analysis Software.

### BRET and PLA Assays

Luminescence and fluorescence signals of COS-1 cells transiently transfected with plasmids encoding Rluc BRET donor and YFP BRET acceptor fusion proteins were analyzed using an Infinite F500 microplate reader (Tecan).

PLA was performed using Duolink reagents (Sigma) per manufacturer's instructions.

### Calcium Imaging and Electrophysiological Recordings

For calcium imaging experiments, primary sensory neurons or HEK293 cells were loaded with the calcium indicators Fura-2 or Cal-520 AM. Electrophysiological recordings were performed using a 700B amplifier and 1440A Analog-Digital Converter (Molecular Devices) in whole-cell voltage clamp configuration.

### GABA Measurements

GABA concentrations were determined by LC-MS/MS at Brainsonline (Groningen).

### Statistical Analyses

Data are presented as mean  $\pm$  SEM. Statistical significance was evaluated employing Student's t test for paired comparisons unless indicated otherwise. p values  $< 0.05$  were considered statistically significant, with \*p < 0.05, \*\*p < 0.01, and \*\*\*p < 0.001.

Additional Information is available in the Extended Experimental Procedures.

## SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, seven figures, one table, and one movie and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2015.01.022>.

## AUTHOR CONTRIBUTIONS

J.S. conceived the project and designed the experiments together with C.H., M.M., W.C.L., and H.W., who performed and analyzed experiments. M.K., L.A., K.S.-S., A.T.-T., C.W., H.K., M.G., and D.R. conducted and analyzed experiments. B.B., M.G., G.R.L., and M.S. contributed to the study design and experiments. J.S. wrote the paper. All authors commented on and approved the paper.

## ACKNOWLEDGMENTS

We thank David Julius, Rohini Kuner, Thomas Müller, Björn Schröder, Christoph Stein, Gerd Zolles, and Bernd Fakler for critical tools and reagents. We also thank Jana Rossius and Christina Steinmeyer-Stannek for expert technical support and Christopher Bohlen for helpful criticism and reading of the manuscript. This work was supported by the European Research Council (ERC-2011-StG-280565 to J.S. and ERC-2011-AdG-294678 to G.R.L.) and the Alexander von Humboldt Foundation/BMBF to M.M. and J.S. Additional support came from the National Center for Competences in Research "Synapsy, Synaptic Bases of Mental Health Disease" Swiss National Science Foundation (3100A0-117816) to B.B. Dr. Roggenkamp is an employee of Beiersdorf AG.

Received: April 22, 2014

Revised: November 16, 2014

Accepted: January 6, 2015

Published: February 12, 2015

## REFERENCES

- Adelfinger, L., Turecek, R., Ivankova, K., Jensen, A.A., Moss, S.J., Gassmann, M., and Bettler, B. (2014). GABA<sub>B</sub> receptor phosphorylation regulates KCTD12-induced K<sup>+</sup> current desensitization. *Biochem. Pharmacol.* 91, 369–379.
- Ayoub, M.A., and Pfleger, K.D. (2010). Recent advances in bioluminescence resonance energy transfer technologies to study GPCR heteromerization. *Curr. Opin. Pharmacol.* 10, 44–52.
- Binet, V., Brajon, C., Le Corre, L., Acher, F., Pin, J.P., and Prézeau, L. (2004). The heptahelical domain of GABA(B2) is activated directly by CGP7930, a positive allosteric modulator of the GABA(B) receptor. *J. Biol. Chem.* 279, 29085–29091.
- Bonnington, J.K., and McNaughton, P.A. (2003). Signalling pathways involved in the sensitisation of mouse nociceptive neurones by nerve growth factor. *J. Physiol.* 557, 433–446.
- Cao, E., Liao, M., Cheng, Y., and Julius, D. (2013). TRPV1 structures in distinct conformations reveal activation mechanisms. *Nature* 504, 113–118.
- Casanova, E., Guetg, N., Vigot, R., Seddik, R., Julio-Pieper, M., Hyland, N.P., Cryan, J.F., Gassmann, M., and Bettler, B. (2009). A mouse model for visualization of GABA(B) receptors. *Genesis* 47, 595–602.
- Caterina, M.J., Schumacher, M.A., Tominaga, M., Rosen, T.A., Levine, J.D., and Julius, D. (1997). The capsaicin receptor: a heat-activated ion channel in the pain pathway. *Nature* 389, 816–824.
- Cavanaugh, D.J., Chesler, A.T., Bráz, J.M., Shah, N.M., Julius, D., and Basbaum, A.I. (2011). Restriction of transient receptor potential vanilloid-1 to the peptidergic subset of primary afferent neurons follows its developmental downregulation in nonpeptidergic neurons. *J. Neurosci.* 31, 10119–10127.
- Charles, K.J., Evans, M.L., Robbins, M.J., Calver, A.R., Leslie, R.A., and Pangalos, M.N. (2001). Comparative immunohistochemical localisation of GABA(B1a), GABA(B1b) and GABA(B2) subunits in rat brain, spinal cord and dorsal root ganglion. *Neuroscience* 106, 447–467.

Chuang, H.H., Prescott, E.D., Kong, H., Shields, S., Jordt, S.E., Basbaum, A.I., Chao, M.V., and Julius, D. (2001). Bradykinin and nerve growth factor release the capsaicin receptor from PtdIns(4,5)P<sub>2</sub>-mediated inhibition. *Nature* 411, 957–962.

Davis, J.B., Gray, J., Gunthorpe, M.J., Hatcher, J.P., Davey, P.T., Overend, P., Harries, M.H., Latcham, J., Clapham, C., Atkinson, K., et al. (2000). Vanilloid receptor-1 is essential for inflammatory thermal hyperalgesia. *Nature* 405, 183–187.

Fischer, M.J., Btesh, J., and McNaughton, P.A. (2013). Disrupting sensitization of transient receptor potential vanilloid subtype 1 inhibits inflammatory hyperalgesia. *J. Neurosci.* 33, 7407–7414.

Forster, A.B., Reeh, P.W., Messlinger, K., and Fischer, M.J. (2009). High concentrations of morphine sensitize and activate mouse dorsal root ganglia via TRPV1 and TRPA1 receptors. *Mol. Pain* 5, 17.

Galvez, T., Duthey, B., Kniazeff, J., Blahos, J., Rovelli, G., Bettler, B., Prézeau, L., and Pin, J.P. (2001). Allosteric interactions between GB1 and GB2 subunits are required for optimal GABA(B) receptor function. *EMBO J.* 20, 2152–2159.

Gassmann, M., Shaban, H., Vigot, R., Sansig, G., Haller, C., Barbieri, S., Humeau, Y., Schuler, V., Müller, M., Kinzel, B., et al. (2004). Redistribution of GABAB(1) protein and atypical GABAB responses in GABAB(2)-deficient mice. *J. Neurosci.* 24, 6086–6097.

Gold, M.S., and Gebhart, G.F. (2010). Nociceptor sensitization in pain pathogenesis. *Nat. Med.* 16, 1248–1257.

Gu, Q., Kwong, K., and Lee, L.Y. (2003). Ca<sup>2+</sup> transient evoked by chemical stimulation is enhanced by PGE2 in vagal sensory neurons: role of cAMP/PKA signaling pathway. *J. Neurophysiol.* 89, 1985–1993.

Haller, C., Casanova, E., Müller, M., Vacher, C.M., Vigot, R., Doll, T., Barbieri, S., Gassmann, M., and Bettler, B. (2004). Floxed allele for conditional inactivation of the GABAB(1) gene. *Genesis* 40, 125–130.

Huang, J., Zhang, X., and McNaughton, P.A. (2006). Inflammatory pain: the cellular basis of heat hyperalgesia. *Curr. Neuropharmacol.* 4, 197–206.

Jones, K.A., Borowsky, B., Tamm, J.A., Craig, D.A., Durkin, M.M., Dai, M., Yao, W.J., Johnson, M., Gunwaldsen, C., Huang, L.Y., et al. (1998). GABA(B) receptors function as a heteromeric assembly of the subunits GABA(B)R1 and GABA(B)R2. *Nature* 396, 674–679.

Kaupmann, K., Malitschek, B., Schuler, V., Heid, J., Froestl, W., Beck, P., Mosbacher, J., Bischoff, S., Kulik, A., Shigemoto, R., et al. (1998). GABA(B)-receptor subtypes assemble into functional heteromeric complexes. *Nature* 396, 683–687.

Kuner, R., Köhr, G., Grünewald, S., Eisenhardt, G., Bach, A., and Kornau, H.C. (1999). Role of heteromer formation in GABA(B) receptor function. *Science* 283, 74–77.

Laffray, S., Tan, K., Dulluc, J., Bouali-Benazzouz, R., Calver, A.R., Nagy, F., and Landry, M. (2007). Dissociation and trafficking of rat GABA(B) receptor heterodimer upon chronic capsaicin stimulation. *Eur. J. Neurosci.* 25, 1402–1416.

Liu, F., Wan, Q., Pristupa, Z.B., Yu, X.M., Wang, Y.T., and Niznik, H.B. (2000). Direct protein-protein coupling enables cross-talk between dopamine D5 and gamma-aminobutyric acid A receptors. *Nature* 403, 274–280.

Loo, L., Shepherd, A.J., Mickle, A.D., Lorca, R.A., Shutov, L.P., Usachev, Y.M., and Mohapatra, D.P. (2012). The C-type natriuretic peptide induces thermal hyperalgesia through a noncanonical Gβγ-dependent modulation of TRPV1 channel. *J. Neurosci.* 32, 11942–11955.

Lopshire, J.C., and Nicol, G.D. (1998). The cAMP transduction cascade mediates the prostaglandin E2 enhancement of the capsaicin-elicited current in rat sensory neurons: whole-cell and single-channel studies. *J. Neurosci.* 18, 6081–6092.

Lu, F.F., Su, P., Liu, F., and Daskalakis, Z.J. (2012). Activation of GABA(B) receptors inhibits protein kinase B/glycogen synthase kinase 3 signaling. *Mol. Brain* 5, 41.

Mandadi, S., Tominaga, T., Numazaki, M., Murayama, N., Saito, N., Armati, P.J., Roufogalis, B.D., and Tominaga, M. (2006). Increased sensitivity of desensitized TRPV1 by PMA occurs through PKCε-mediated phosphorylation at S800. *Pain* 123, 106–116.

Margeta-Mitrovic, M., Jan, Y.N., and Jan, L.Y. (2000). A trafficking checkpoint controls GABA(B) receptor heterodimerization. *Neuron* 27, 97–106.

Miller, K.E., Hoffman, E.M., Sutharshan, M., and Schechter, R. (2011). Glutamate pharmacology and metabolism in peripheral primary afferents: physiological and pathophysiological mechanisms. *Pharmacol. Ther.* 130, 283–309.

Moriyama, T., Higashi, T., Togashi, K., Iida, T., Segi, E., Sugimoto, Y., Tomi-naga, T., Narumiya, S., and Tominaga, M. (2005). Sensitization of TRPV1 by EP1 and IP reveals peripheral nociceptive mechanism of prostaglandins. *Mol. Pain* 1, 3.

Padgett, C.L., and Slesinger, P.A. (2010). GABA(B) receptor coupling to G-proteins and ion channels. *Adv. Pharmacol.* 58, 123–147.

Pagano, A., Rovelli, G., Mosbacher, J., Lohmann, T., Duthey, B., Stauffer, D., Ristig, D., Schuler, V., Meigel, I., Lampert, C., et al. (2001). C-terminal interaction is essential for surface trafficking but not for heteromeric assembly of GABA(B) receptors. *J. Neurosci.* 21, 1189–1202.

Perroy, J., Adam, L., Qanbar, R., Chénier, S., and Bouvier, M. (2003). Phosphorylation-independent desensitization of GABA(B) receptor by GRK4. *EMBO J.* 22, 3816–3824.

Por, E.D., Bierbower, S.M., Berg, K.A., Gomez, R., Akopian, A.N., Wetzel, W.C., and Jeske, N.A. (2012). β-Arrestin-2 desensitizes the transient receptor potential vanilloid 1 (TRPV1) channel. *J. Biol. Chem.* 287, 37552–37563.

Rutter, A.R., Ma, Q.P., Leveridge, M., and Bonnert, T.P. (2005). Heteromerization and colocalization of TrpV1 and TrpV2 in mammalian cell lines and rat dorsal root ganglia. *Neuroreport* 16, 1735–1739.

Schwenk, J., Metz, M., Zolles, G., Turecek, R., Fritzius, T., Bildl, W., Tarusawa, E., Kulik, A., Unger, A., Ivankova, K., et al. (2010). Native GABA(B) receptors are heterooligomers with a family of auxiliary subunits. *Nature* 465, 231–235.

Siemens, J., Zhou, S., Piskorowski, R., Nikai, T., Lumpkin, E.A., Basbaum, A.I., King, D., and Julius, D. (2006). Spider toxins activate the capsaicin receptor to produce inflammatory pain. *Nature* 444, 208–212.

Söderberg, O., Gullberg, M., Jarvius, M., Ridderstråle, K., Leuchowius, K.J., Jarvius, J., Wester, K., Hydbring, P., Bahram, F., Larsson, L.G., and Landegren, U. (2006). Direct observation of individual endogenous protein complexes in situ by proximity ligation. *Nat. Methods* 3, 995–1000.

Sokal, D.M., and Chapman, V. (2003). Inhibitory effects of spinal baclofen on spinal dorsal horn neurones in inflamed and neuropathic rats in vivo. *Brain Res.* 987, 67–75.

Sudo, Y., Hojo, M., Ando, Y., Takada, M., Murata, H., Kurata, S., Nishida, N., and Uezono, Y. (2012). GABA(B) receptors do not internalize after baclofen treatment, possibly due to a lack of β-arrestin association: study with a real-time visualizing assay. *Synapse* 66, 759–769.

Szallasi, A., and Sheta, M. (2012). Targeting TRPV1 for pain relief: limits, losers and laurels. *Expert Opin. Investig. Drugs* 21, 1351–1369.

Tominaga, M., Caterina, M.J., Malmberg, A.B., Rosen, T.A., Gilbert, H., Skinner, K., Raumann, B.E., Basbaum, A.I., and Julius, D. (1998). The cloned capsaicin receptor integrates multiple pain-producing stimuli. *Neuron* 21, 531–543.

Towers, S., Princivalle, A., Billinton, A., Edmunds, M., Bettler, B., Urban, L., Castro-Lopes, J., and Bowery, N.G. (2000). GABA(B) receptor protein and mRNA distribution in rat spinal cord and dorsal root ganglia. *Eur. J. Neurosci.* 12, 3201–3210.

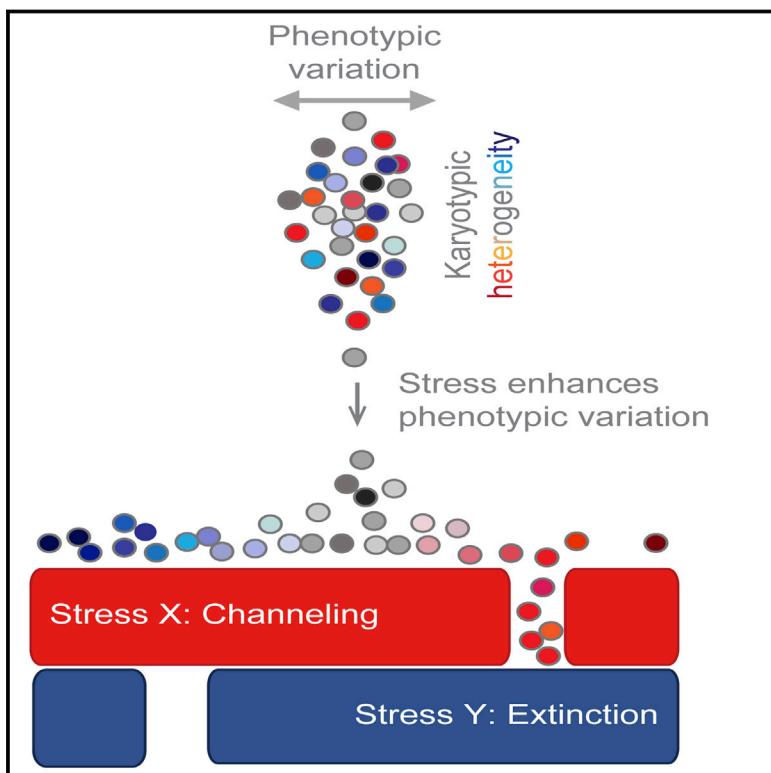
Vay, L., Gu, C., and McNaughton, P.A. (2012). The thermo-TRP ion channel family: properties and therapeutic implications. *Br. J. Pharmacol.* 165, 787–801.

White, J.H., Wise, A., Main, M.J., Green, A., Fraser, N.J., Disney, G.H., Barnes, A.A., Emson, P., Foord, S.M., and Marshall, F.H. (1998). Heterodimerization is required for the formation of a functional GABA(B) receptor. *Nature* 396, 679–682.

Woolf, C.J. (2010). Overcoming obstacles to developing new analgesics. *Nat. Med.* 16, 1241–1247.

# Targeting the Adaptability of Heterogeneous Aneuploids

## Graphical Abstract



## Authors

Guangbo Chen, Wahid A. Mulla, ...,  
Judith Berman, Rong Li

## Correspondence

rli@stowers.org

## In Brief

The heterogeneity of aneuploid cell populations increases with stress, causing resistance to emerge. An eradication strategy involves sequential applications of stress: the first stress homogenizes the population via adaptation, and the second specifically targets and eliminates the newly dominant karyotype.

## Highlights

- Higher stress leads to larger phenotypic variation in heterogeneous aneuploids
- Eradication of aneuploids via dual-stress application: “evolutionary trap” (ET)
- The first stress selects for a less diverse population; the second targets it
- ET may be applicable toward azole resistance in *Candida albicans* and human cancer

# Targeting the Adaptability of Heterogeneous Aneuploids

Guangbo Chen,<sup>1,2</sup> Wahid A. Mulla,<sup>1,2</sup> Andrei Kucharavy,<sup>1,3</sup> Hung-Ji Tsai,<sup>1</sup> Boris Rubinstein,<sup>1</sup> Juliana Conkright,<sup>1</sup> Scott McCroskey,<sup>1</sup> William D. Bradford,<sup>1</sup> Lauren Weems,<sup>1</sup> Jeff S. Haug,<sup>1</sup> Chris W. Seidel,<sup>1</sup> Judith Berman,<sup>4</sup> and Rong Li<sup>1,2,\*</sup>

<sup>1</sup>Stowers Institute for Medical Research, 1000 East 50<sup>th</sup> Street, Kansas City, MO 64110, USA

<sup>2</sup>Department of Molecular and Integrative Physiology, University of Kansas Medical Center, 3901 Rainbow Boulevard, Kansas City, KS 66160, USA

<sup>3</sup>Sorbonne Universités, UPMC Univ Paris 06, UMR 7238, Biologie Computationnelle et Quantitative, F-75005 Paris, France

<sup>4</sup>Department of Molecular Microbiology and Biotechnology, George Wise Faculty of Life Sciences, Tel Aviv University, Ramat Aviv 69978, Israel

\*Correspondence: rli@stowers.org

<http://dx.doi.org/10.1016/j.cell.2015.01.026>

## SUMMARY

Aneuploid genomes, characterized by unbalanced chromosome stoichiometry (karyotype), are associated with cancer malignancy and drug resistance of pathogenic fungi. The phenotypic diversity resulting from karyotypic diversity endows the cell population with superior adaptability. We show here, using a combination of experimental data and a general stochastic model, that the degree of phenotypic variation, thus evolvability, escalates with the degree of overall growth suppression. Such scaling likely explains the challenge of treating aneuploidy diseases with a single stress-inducing agent. Instead, we propose the design of an “evolutionary trap” (ET) targeting both karyotypic diversity and fitness. This strategy entails a selective condition “channeling” a karyotypically divergent population into one with a predominant and predictably drugable karyotypic feature. We provide a proof-of-principle case in budding yeast and demonstrate the potential efficacy of this strategy toward aneuploidy-based azole resistance in *Candida albicans*. By analyzing existing pharmacogenomics data, we propose the potential design of an ET against glioblastoma.

## INTRODUCTION

Germline evolution shapes the organismal tree in changing environments during the long course of natural history, whereas acute environmental fluctuation drives asexual cellular evolution, often associated with dynamic structural changes in the genomes of microbes or cancer cells (Lewontin, 1970; Merlo et al., 2006). Aneuploidy (chromosome copy-number imbalance) is a type of genome alteration widely observed during cellular evolution of eukaryotic species, such as laboratory (Chen et al., 2012a; Hughes et al., 2000; Rancati et al., 2008), industrial (Borneman et al., 2011; Chen et al., 2012b; Infante et al., 2003;

Kvitek et al., 2008), and pathogenic (Marichal et al., 1997; Ni et al., 2013; Selmecki et al., 2006; Sionov et al., 2010) yeasts, as well as protozoan parasites such as leishmania (Leprohon et al., 2009; Mannaert et al., 2012; Ubeda et al., 2008) and trypanosomes (Llewellyn et al., 2011; Minning et al., 2011). Emerging evidence also points to aneuploidy as an important driver for the evolution of human cancer (Davoli et al., 2013; Holland and Cleveland, 2009; Jones et al., 2010; Ng et al., 2010; Wang et al., 2014). Due to its impact on the expression of many genes, aneuploidy brings about large phenotypic changes that can be either detrimental or beneficial in a karyotype- and environmental condition-dependent manner (Chen et al., 2012b; Pavelka et al., 2010).

Aneuploid populations are often characterized by heterogeneity—the coexistence of many different karyotypes. The genetic diversity provides the raw material for evolutionary selection and endows the aneuploid population with high adaptive potential (Burrell et al., 2013; Chen et al., 2012a; Maley et al., 2006). This underscores the exceptional challenge of treating disease-causing cell populations characterized by large karyotype heterogeneity and instability (Gerlinger et al., 2012; Harrison et al., 2014; Lee et al., 2011; Navin et al., 2011; Sotillo et al., 2010). One idea is to find drugs strongly exacerbating a common deficiency of aneuploids irrespective of specific karyotype (Orromendia and Amon, 2014). A study found that among a collection of disomic yeast strains, many showed prominent growth defects toward agents that perturb proteome homeostasis, such as hygromycin B (a translation inhibitor) or geldanamycin (an Hsp90 inhibitor) (Torres et al., 2007), supporting the notion that a common deficiency of aneuploidy is overloading of the proteome quality-control system. However, some of the disomic strains exhibited resistance, as opposed to sensitivity, toward proteotoxic agents, and one of the karyotypic features, increased copy number of chromosome XV (chrXV), emerged as the adaptive variant when a diploid strain was evolved in the presence of an Hsp90 inhibitor (Chen et al., 2012a). In addition, when a highly heterogeneous aneuploid population was treated with drugs that imposed strong immediate growth inhibition, long-term culturing enabled evolutionary selection for a few variants that eventually rendered the population drug resistant (Chen et al., 2012a) (also see results of this study). These findings

highlight the fallacy of short-term efficacy in drug treatments when dealing with heterogeneous populations already poised for rapid adaptation.

In this study, we aim for an innovative approach that accounts for the evolutionary dynamics and achieves long-term growth suppression or extinction of aneuploid cell populations consisting of a wide spectrum of karyotypes. Our analyses demonstrate on a general level that the evolutionary potential of a heterogeneous aneuploid population escalates under increasing stress, but such a population may become highly targetable once the karyotype diversity is drastically confined. These findings led us to design a two-component strategy for treating diseases associated with aneuploid cell populations that targets the population's adaptability and fitness.

## RESULTS

### Phenotypic Variation Scales with Growth Suppression under Diverse Stress Conditions in Budding Yeast

To investigate whether certain stress conditions may be consistently effective toward aneuploids with a wide-spectrum of karyotypes, we performed a re-analysis of the growth data from a previous study subjecting a panel of 38 aneuploid yeast strains (*S. cerevisiae*) with diverse and random chromosome stoichiometry to phenotypic profiling across a wide range of conditions with varying stress types and levels (Pavelka et al., 2010) (Figure 1A). This analysis revealed a surprising trend: those conditions more toxic to the aneuploidy cohort (lower mean growth ratio) also produced a larger fitness spread among the aneuploids, such that even under highly toxic conditions, although many aneuploids failed to grow, some aneuploids endured or thrived. The standard deviation (spread) and mean of the growth ratio (overall growth suppression) of diverse aneuploids are strongly correlated across diverse stress conditions (Figure 1B). This same trend could also be observed with the set of disomic aneuploid strains isolated in a previous study (Torres et al., 2007) treated with increasing concentrations of different drugs (Figure 1C).

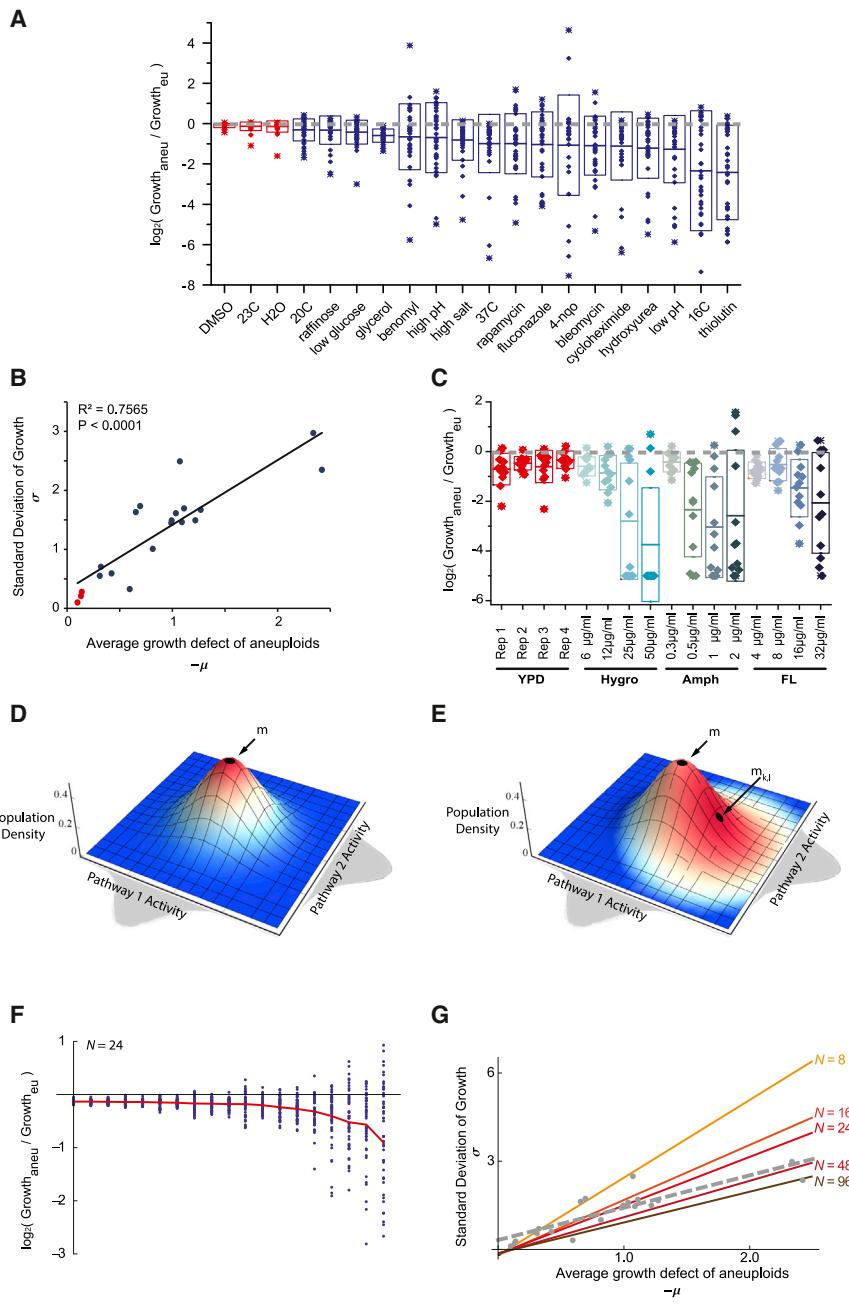
We used a stochastic model to investigate the theoretical generality of the above experimental observation (Figures 1D, 1E, S1A, and S1B; see [Extended Experimental Procedures](#) for a detailed description of assumptions and simulations). Briefly, we assumed that fitness is governed by  $N$  independent pathways, reflecting the modular architecture of cellular systems. We assume that for each individual pathway the distribution of pathway activity across diverse karyotypes assumes a simple normal distribution. Under the stress-free condition, the optimal fitness is reached by the euploid at the peak of the  $N$ -dimensional normal distribution (Figure 1D; [Extended Experimental Procedures](#)), whereas aneuploid populations are located at positions away from the peak, reflecting suboptimal fitness. In the presence of stress, however, the optimal fitness point moves away from the euploid position (Figure 1E). The change in the distance between the position of each karyotype and the optimal fitness point represents the fitness change caused by stress. Simulation of this model with varying values of  $N$  revealed two properties of the system remarkably consistent with experimental observations (Figures 1F and 1G). First, under stress con-

ditions with sufficient magnitude, most aneuploids had lower fitness than the euploid, yet there was always a fraction of aneuploids that assumed higher fitness than the euploid. Second, the absolute value of the mean and standard deviation of the relative fitness satisfied a positive linear correlation across diverse stress conditions and a range of  $N$  values (Figures 1F, 1G, and S1C–S1F). The fact that such a simple model could qualitatively capture the experimental data suggests the observed phenomenon to be a general property of heterogeneous aneuploid populations irrespective of specific stress applied.

### Therapeutic Compounds Elicit Phenotypic Heterogeneity in Human Tumor Cell Lines

A vast majority of solid tumors are aneuploid: 91.7% of 817 solid tumor cell lines surveyed by Cancer Cell Line Encyclopedia (CCLE) have at least 1 chromosome arm-level copy-number variation, whereas 58.9% have more than 10 (see [Extended Experimental Procedures](#)). Karyotypic heterogeneity has also been observed within a single tumor (Gerlinger et al., 2012; Navin et al., 2011). The relapse fueled by adaptive changes in a tumor remains a major challenge in cancer treatment. A logical extension of our findings is that, in karyotypically heterogeneous cancer cell populations, the treatment suppressing the overall growth may also escalate tumor cells' phenotypic heterogeneity and potentially adaptability. To examine this possibility, we first investigated a dataset consisting of 54 karyotypically divergent human breast cancer cell lines and 77 different anticancer drugs across 10 concentrations (SU2C breast cancer project; Heiser et al., 2012). As shown in Figures 2A and 2B, in drug doses that caused prominent (>50%) overall growth suppression, the spread of growth rates among cell lines treated with most chemicals (~95%) also increased. A few drugs showed more consistent toxicity, but due to the lack of a normal euploid control, it is unknown whether this was specific to aneuploid cancer cells. Examination of the correlation between overall growth suppression (-mean) and variation (SD of growth rates) found a consistent positive correlation. Simulation of our multi-dimensional fitness model with varying  $N$  values produced a similar correlation pattern (Figures 2C and 2D). Cell lines from other solid tumors (e.g., central nervous system, skin, and lung; Barretina et al., 2012) also demonstrated a similar trend (Figure S2).

Although model simulations showed that for a given stress (X), a group of aneuploids may be selected for superior fitness, they also predict that if a second stress (Y) is applied to shift the optimal fitness in the direction opposite from stress X, those selected aneuploids would fall into the death zone in the fitness landscape (Figures 3A and 3B). This implies that if a highly evolvable population with heterogeneous random karyotypes can be "channeled" toward a certain karyotypic characteristic under a designed selection (stress X), thus drastically "shrinking" the population's evolvability, a second treatment (stress Y) may be added to eradicate the selected singular or limited karyotypes. Together these two treatments would form an "evolutionary trap" (ET) on the elusive aneuploid population. Supporting this idea, for a given specific aneuploid yeast strain, in most cases one or multiple conditions could be found that caused >80% growth inhibition compared to the euploid control (Figure S1G).



**Figure 1. Scaling of Phenotypic Variation with Growth Suppression in Aneuploids under Diverse Stress Conditions**

(A) The growth of 38 aneuploid strains relative to the euploid, as  $\log_2$  ratio of aneuploid growth (OD increase) over the euploid with the nearest ploidy (see [Extended Experimental Procedures](#)), is binned by growth conditions. Each point in a boxplot represents an aneuploid strain. The half-length of each box represents the SD of relative growth among aneuploids ( $\sigma$ ), and the middle line represents the average ( $\mu$ ). Note that the horizontal dashed line across 0 represents the euploid control.

(B) Phenotypic variation among the aneuploids, measured as SD of relative growth ( $\sigma$ ), scales with average growth defect of the aneuploid cohort across diverse stress conditions ( $-\mu$ ).

(C) The growth of 12 disomy strains relative to the haploid control under increasing concentrations of hygromycin B (Hygro), amphotericin B (Amph), or fluconazole (FL). Boxplot representation is as described for (A).

(D) and (E) Schematic representation of the model is shown for the simple case of  $N = 2$  with axes as labeled. Deep blue to deep red code for increasing fitness. (D) Graph represents the stress-free condition, where the euploid that is located at the center of the activity field (position  $m$ ) assumes the highest fitness. (E) Graph represents a stress condition, where the optimal fitness point shifts from  $m$  to  $m_{k,l}$ , reflecting the activity change necessary for adaptation. Consequently, the euploid (located at point  $m$ ) no longer holds maximal fitness, whereas higher fitness is assumed by certain aneuploids (those occupying redder regions).

(F) Example simulation results of the model for 50 random aneuploids under diverse stress conditions (governed by varying type  $k$  and magnitude  $l$ ) for a 24 dimension space ( $N = 24$ ), with relative growth displayed as the experimental data in (A). The red line shows average  $\log_2$  growth ratio from the simulated aneuploid population. Note the appearance of adaptive aneuploids under high-stress conditions (toward the right of the graph).

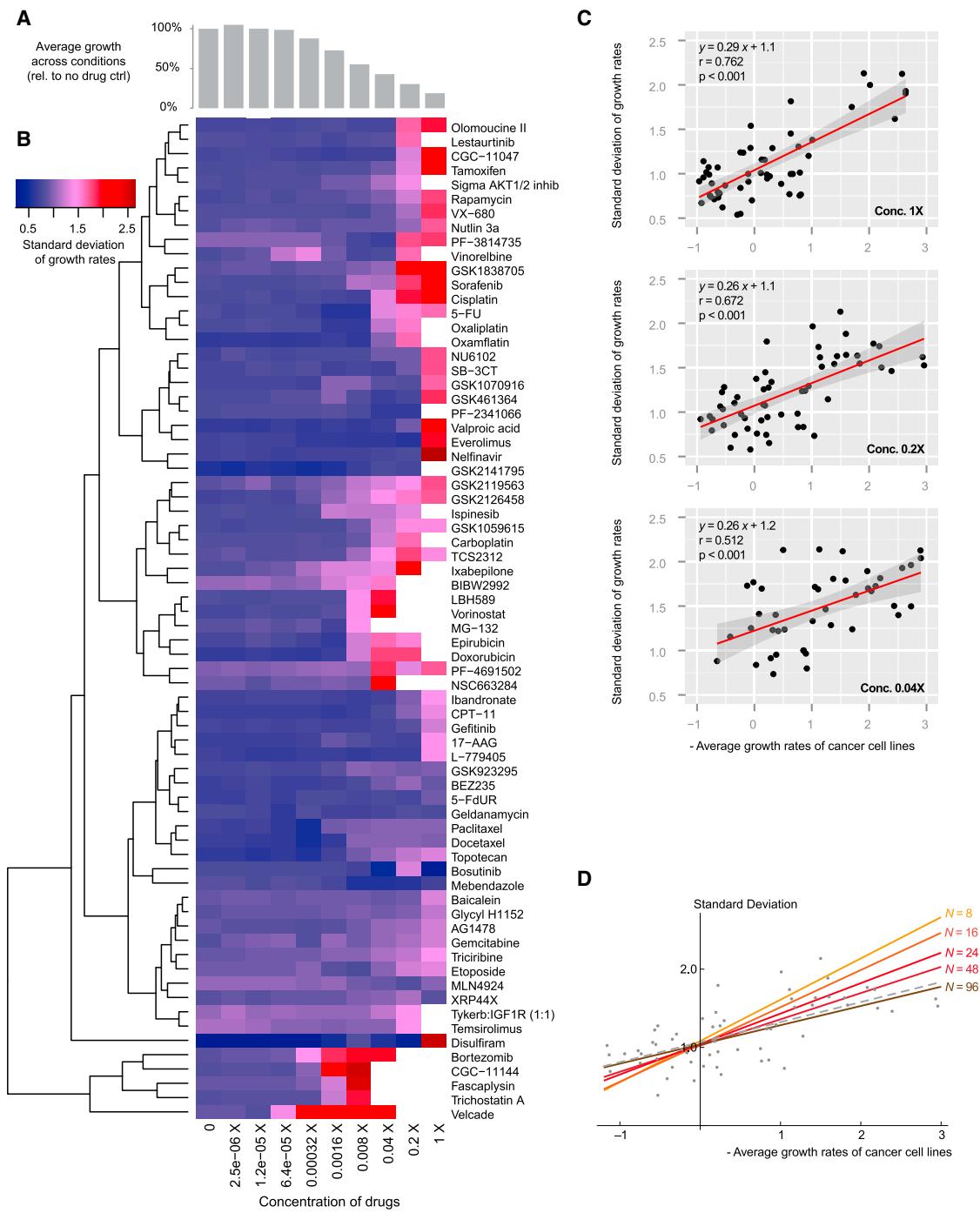
(G) Simulations of the model with a wide range of  $N$  values demonstrate the positive correlation between  $\sigma$  and  $-\mu$  in various numbers of dimensions. The simulated correlations are shown in colored lines, whereas the experimental data are overlaid in gray.

See also [Figure S1](#).

### Finding Drugs Forming an ET against Heterogeneous Aneuploid Budding Yeast

To investigate the validity of ET against heterogeneous aneuploids, we first constructed a yeast population with high-degree karyotypic diversity by sporulation of a homozygous pentaploid yeast strain. Twenty-eight viable aneuploid meiotic products with ploidy above 1.9N were mixed with isogenic diploid, triploid, and tetraploid cells to mimic karyotype heterogeneity observed in pathogenic fungi ([Harrison et al., 2014](#)) or human cancers ([Mitelman et al., 2012](#)). Because most aneuploid karyotypes are unstable ([St Charles et al., 2010](#); [Pavelka et al., 2010](#); [Zhu et al.,](#)

[2012](#)), the degree of heterogeneity of this mixed population (heterogeneous mix) is expected to be far greater, as confirmed by a broad DNA content profile that did not show bias toward any specific chromosome gain or loss ([Figure S3](#)). To find a selective condition for karyotype channeling, we took advantage of the previous finding that growth in high concentrations of radicicol, an inhibitor of Hsp90 chaperone, selects for extra chromosome (chr) XV ([Chen et al., 2012a](#)). Treatment of three independent populations of the heterogeneous mix with 50  $\mu\text{g/ml}$  radicicol all predictably led to convergent evolution toward chrXV gain ([Figures S3C–S3E](#)). Six individual colonies derived from a culture



**Figure 2. Phenotypic Heterogeneity of Human Cancer Cell Lines Positively Correlates with Increasing Growth Suppression by Therapeutic Compounds**

The relationship between phenotypic variation (SD of growth rates) and average growth suppression was examined using published drug-response profiling data from 54 breast cancer cell lines treated with 77 different potential therapeutic compounds over 10 different concentrations (see [Extended Experimental Procedures](#)).

(A) Histograms showing average growth suppression under different drug concentrations.

(B) The SD of growth rates caused by each drug under the range of concentrations tested are shown as a heatmap. Note that drug concentrations for (A) and (B) are aligned, showing the general trend of increasing SD with increasing growth suppression. The clustering is based on Euclidean distances.

(legend continued on next page)

that had grown to saturation in radicicol-containing medium were found to vary in karyotypes but all shared chrXV gain, confirming that radicicol caused a selective sweep in the heterogeneous mix (Figure S3F).

We next performed a screen for chemicals particularly effective against chrXV trisomy and identified hygromycin B, a translation inhibitor (Singh et al., 1979), as the most potent inhibitor of chrXV trisomy relative to its effect on the diploid control (Figure S4). We further confirmed potent growth inhibition by hygromycin B for all six different karyotypes identified from radicicol-selected cultures sharing chrXV gain (Figure 3C). We note that even though hygromycin B effectively suppressed the growth of aneuploids that had gained chrXV, this effect did not extend to all aneuploids: aneuploid strains with increased chrII or IX dosage but having a normal dosage of chrXV were associated with superior resistance to hygromycin B compared to the euploid control (Figures 3D and 3E). However, strains with combinations of chrXV, II, and IX gains were highly sensitive rather than resistant to hygromycin B (Figure 3E).

We showed previously that dosage increase of *ST11* and *PDR5*, two genes present on chrXV, are “driver mutations” underlying chrXV gain-associated radicicol resistance (Chen et al., 2012a). However, hygromycin B sensitivity was unrelated to dosage alterations of these two genes (Figures 4A and 4B). To identify genes on chrXV whose increased copy number could cause hypersensitivity to hygromycin B, we screened each of the 453 genes located on chrXV with its original promoter and carried on a low-copy (centromeric) plasmid (Ho et al., 2009) and identified 5 genes (*CRS5*, *RPS15*, *TRM11*, *RRP6*, *SER1*) (Figures 4C, S5A, and S5B). Integration of one copy of each top three hit (*CRS5*, *RPS15*, *TRM11*) recapitulated up to 50% of the hygromycin B sensitivity of chrXV trisomy in a diploid background (Figure 4E). However, the effects of these genes were not additive (data not shown), suggesting that the hygromycin B sensitivity associated with chrXV gain is an emergent property of gaining many genes carried on this chromosome beyond the three found in the screen. Consistent with this notion, deletion of a copy of each of the three hits (*CRS5*, *RPS15*, *TRM11*) individually did not change hygromycin hypersensitivity of chrXV trisomy (Figure 4F).

Hygromycin B and other proteotoxic agents were previously proposed to enhance the protein quality-control deficit due to increased expression of a large number of genes (Oromendia et al., 2012; Torres et al., 2007). However, the genes whose dosage increase gave rise to hygromycin B sensitivity were not all highly expressed genes compared to the other genes located on chrXV in either euploid or aneuploid with chrXV gain ( $p = 0.56$  for transcripts,  $p = 0.43$  for proteins, Mann-Whitney U test) (Figures 4C, 4D, and S5C). *RPS15*, the only highly expressed gene among the hits, encodes one of 30 proteins within the small ribosomal subunit to which hygromycin B binds (Borovinskaya et al., 2008), suggesting that a direct gene-specific cause underlies

some of the drug sensitivity. Taken together, different molecular mechanisms are likely to account for radicicol resistance and hygromycin hypersensitivity of aneuploid strains with chrXV gain.

### Evolution Dynamics of Cell Populations under Single-Drug Treatment versus ET

Even though hygromycin B alone potently suppressed the growth of the chrXV trisomy strain initially, the treated cell population eventually adapted, as indicated by growth takeoff after ~50 hr (Figure 5A). Karyotyping of the adapted chrXV trisomy culture revealed that the population was a mixture of euploid and aneuploid cells that no longer had chrXV gain, but some now carried an additional copy of chrIX (Figures 5B and 5C), consistent with the result shown in Figures 3D and 3E. This observation reconfirmed the requirement for chrXV gain in hygromycin B sensitivity and demonstrated the ability of the population to escape single-drug treatment by continued karyotype change. As expected, the condition that selected for chrXV, i.e., 50 µg/ml radicicol, was highly toxic to all those survivors that had become hygromycin B resistant because they had lost the gained chrXV (Figures 5D and 5E), supporting the rationale of combinatorial treatment with both radicicol and hygromycin B.

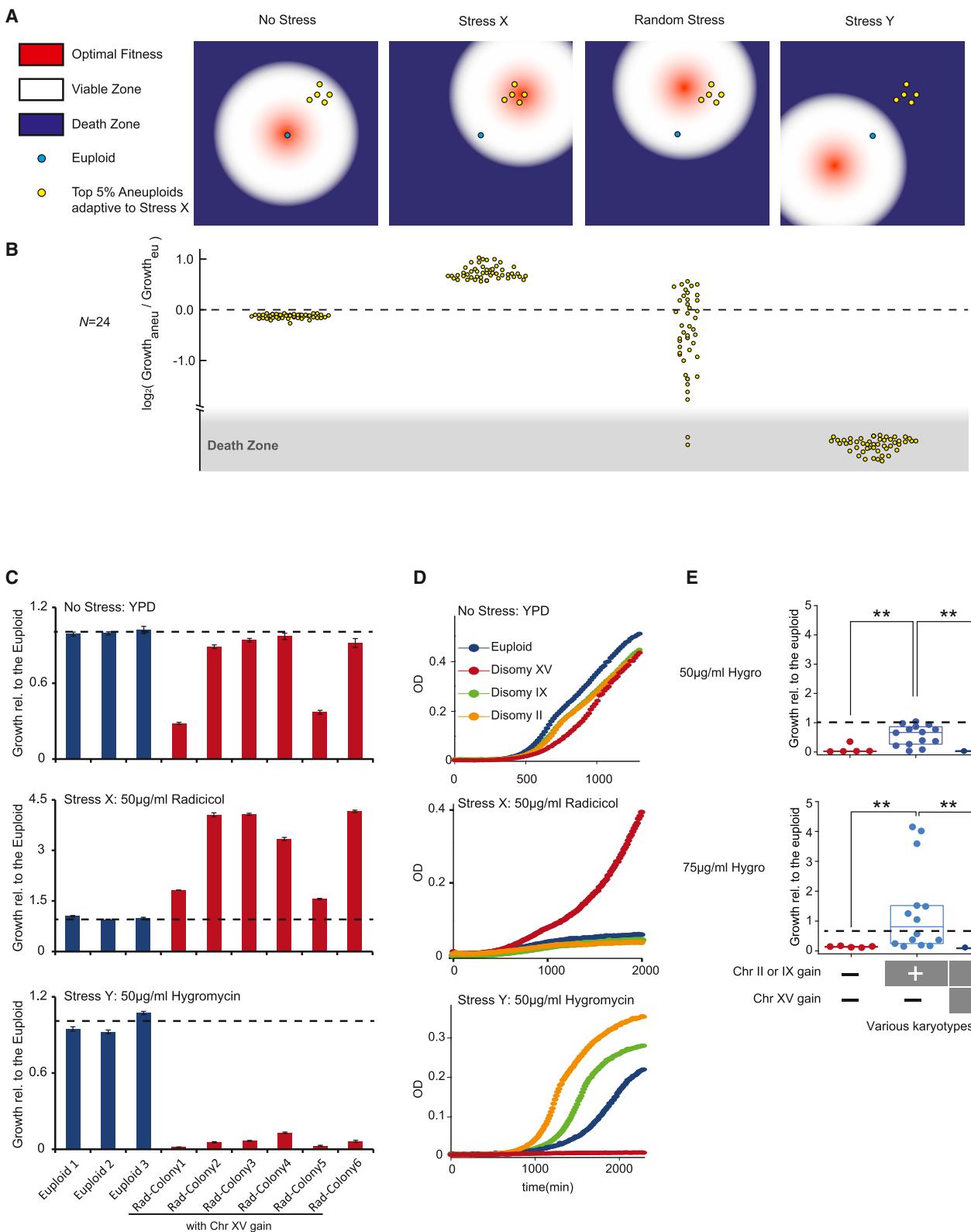
Indeed, the combination of both drugs led to the extinction of all three independent cultures first subjected to the radicicol selection (Figures 5F–5H). The same drug pair was similarly effective against the heterogeneous aneuploid population (heterogeneous mix) when added simultaneously (Figures S6A–S6E). When the heterogeneous mix was treated with a single chemical, 10 out of 21 tested chemicals (for example, 0.08 µg/ml menadoine and 100 µg/ml radicicol) imposed stronger growth suppression than 50 µg/ml hygromycin B (Figure S6C). Hygromycin B also ranked low in suppressing euploid growth (Figure S6A). Yet, when combined with 50 µg/ml radicicol, hygromycin B, but not menadoine or radicicol, led to the extinction of the heterogeneous mix (Figure S6D). Thus, the opposing selective effects of these two drugs on the channeled karyotype (chrXV gain) imposed an adaptive dilemma for the heterogeneous population, leading to its extinction. A potential pitfall of this ET drug pair is that the diploid with an additional copy of *ST11* showed an enhanced level of resistance against 50 µg/ml radicicol yet without causing hypersensitivity to 50 µg/ml hygromycin (Figure 4A). Such effects due to potential single-gene gain could be remedied by increasing the concentration (from 50 to 100 µg/ml) of radicicol (Figures S6F and S6G).

### An Azole-Based ET for Human Pathogen *Candida albicans*

We tested the possibility of using the ET strategy for more effective anti-fungal treatment. It was previously shown that a mechanism for human pathogen *Candida albicans* to confer resistance to fluconazole, a first-line medicine treating invasive

(C) At three concentrations that considerably reduced the overall growth rate (>50% decrease compared to no drug control in A), the general correlation between  $\mu$  and  $\sigma$  across different growth conditions is examined. The linear regression line (red) is surrounded by 95 percentile confidence fitting intervals (darker gray area). Note similar fitting parameters across different drug doses.

(D) The correlation between SD and average growth suppression is also recapitulated by simulations of the multi-dimensional model with the number of pathways ( $N$ ) in the range of 48–96. The modeled fitting is shown in colored lines, whereas the published experimental data are shown in gray. See also Figure S2.



(legend on next page)

candidiasis in immune-compromised patients, is gain of isochromosome 5L (i(5L), which contains two copies of the left arm of chr5) (Selmecki et al., 2006). This aneuploid feature was also recapitulated in a laboratory evolution experiment selecting for fluconazole-resistant *C. albicans* (Selmecki et al., 2009). Thus, fluconazole may serve as the “selection” drug, which would be used in combination with a second drug specifically targeting i(5L) gain. To this end, we screened a chemical library that contained 1,713 FDA- or other regulatory agency-approved drugs, 580 natural compounds, and 420 other bioactive agents against an i(5L)-containing *C. albicans* isolate from a 30-year-old male who developed fluconazole resistance (Marr et al., 1997; Selmecki et al., 2006). We looked for chemicals that showed much elevated potency against the i(5L)-containing strain compared to the euploid control generated by spontaneous i(5L) loss of the original i(5L)-containing strain (Selmecki et al., 2008). Such drugs were likely to be missed in previous screens against euploid *C. albicans* (Okoli et al., 2009; Spitzer et al., 2011).

A primary screen found that 100 out of 2,713 compounds (Figure S7A) caused at least 80% growth suppression to either the i(5L) strain or the euploid, or both. For this set of compounds, we then determined the concentration causing 80% growth suppression ( $IC_{80}$ ). In the presence of 26 or 11 conditions, the i(5L) strain showed significantly (Z test,  $p < 0.05$ ) higher (i.e., more resistant) or lower (more susceptible), respectively,  $IC_{80}$  than the euploid (Figure 6A). As expected, the i(5L) strain exhibited increased resistance toward a panel of seven other azole derivatives. By contrast, pyrvium pamoate (PP), a medication for pediatric pinworm infection, strongly suppressed the growth of the i(5L) strain but was ineffective against the euploid even at high concentrations as high as 80  $\mu$ M after 48 hr culture on agar (Figures 6A, 6B, and S7B). Given that euploid and i(5L)-gained *C. albicans* could co-exist in patients treated with fluconazole (Selmecki et al., 2008), we tested the combinatorial effect of PP with fluconazole in a 1:1 mixed population of both karyotypes. At concentrations of PP that suppressed the growth of the i(5L) strain, the resistance to fluconazole regressed from over 256  $\mu$ g/ml back to the same level exhibited by the euploid without PP (Figures 6C and 6D), supporting the ET rationale for combinatorial anti-fungal treatment.

### A Potential ET against EGFR-Driven Tumors such as Glioblastoma

In many human cancers, epidermal growth factor receptor (EGFR) signaling is a major contributing factor to aggressive phenotypes (Wheeler et al., 2010). The human EGFR gene is located on the short arm of chr7 (chr7p), and chr7p gain is a prevalent karyotypic feature across diverse types of human cancers (Beroukhim et al., 2010). In particular, chr7 gain was observed in 80% of glioblastoma samples collected from The Cancer Genome Atlas Pilot Project (TCGA) cohort of 219 patients (Bredel et al., 2009; Network, 2008). Within individual tumors, arm-level (mainly chr7p) or whole-chromosome gain of chr7p was detected in nearly all surgical sections examined, suggesting that this event, unlike other heterogeneous genetic alterations, is an early critical event in clonal expansion (Sottoriva et al., 2013). In addition to being a likely product of the selection for EGFR-dependent malignant transformation or adaptation, chr7p polysomy was also reported in 80% of tumor samples from lung cancer patients that acquired resistance to gefitinib or erlotinib, EGFR kinase inhibitors (Bean et al., 2007). Thus, in designing an ET against EGFR signaling-dependent tumors, gefitinib or erlotinib may be used as the channeling drug that could impose a pharmacological selection for chr7p gain beyond the selection during tumorigenesis. To find the pairing drug that may be particularly toxic toward chr7p gain, we analyzed the pharmacogenomics data (Barretina et al., 2012) by correlating the dosage of chr7p of 29 central nervous system (CNS) tumor cell lines with the response ( $IC_{50}$ ) to 23 therapeutic compounds (Figure 7A). This analysis identified a significant correlation between increased chr7p dosage and increased sensitivity to irinotecan, a drug that is FDA approved for treating colon cancer but not yet for glioblastoma (Figure 7B). This result suggests that gefitinib or erlotinib together with irinotecan may form an ET against EGFR-driven tumors, especially glioblastoma where chr7p is already highly enriched even without drug treatment (Figure 7E).

### DISCUSSION

The analyses presented above started from an exploration of the general response of karyotypically heterogeneous cell populations to diverse stress conditions and proceeded to the design

#### Figure 3. Design and Experimental Implementation of ET in Budding Yeast

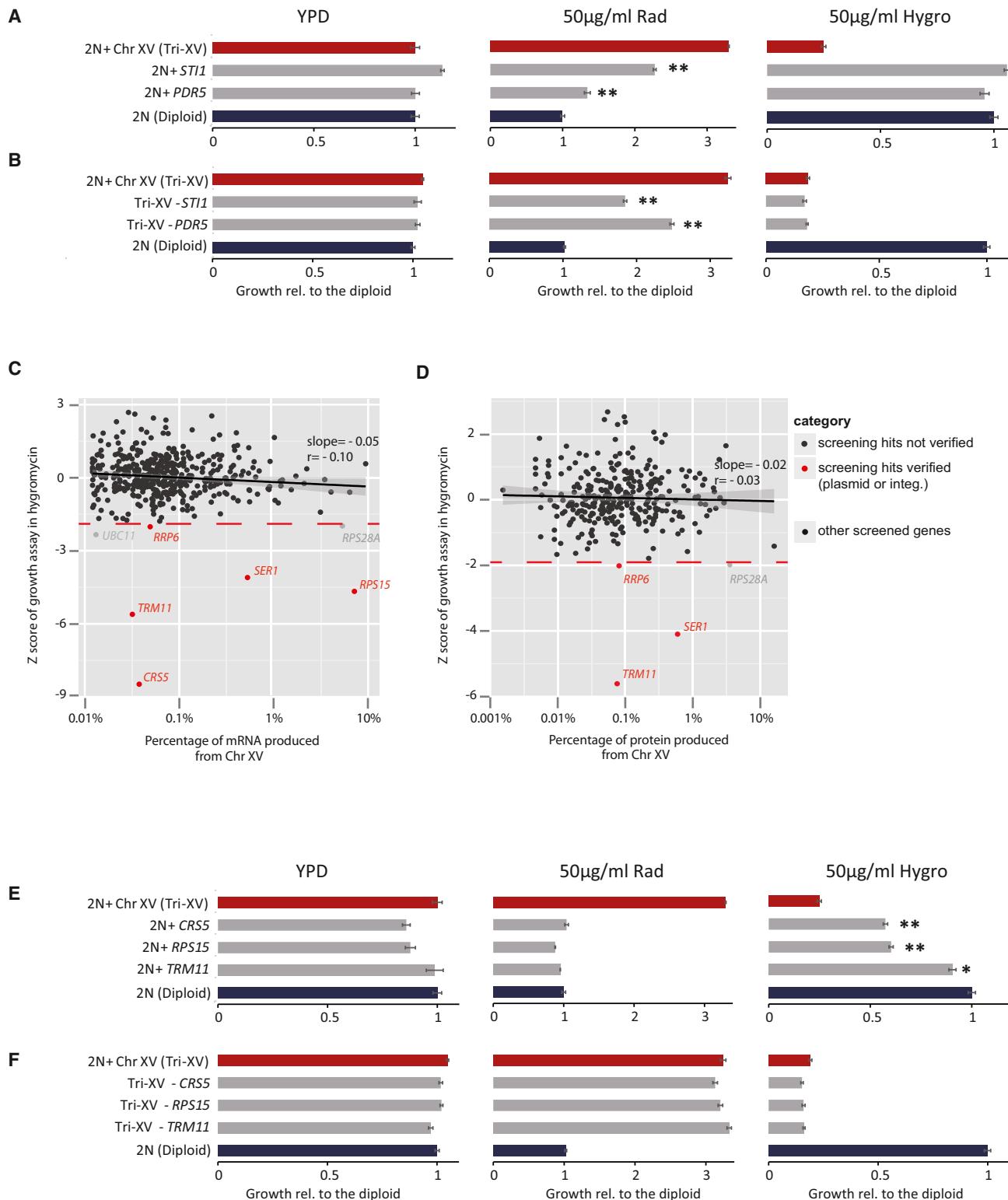
(A and B) Model simulations predict that cell population adapted to a specific stress (X) through karyotype channeling can be highly targetable by a stress Y that shifts optimal fitness in the direction opposite to X, but not by a second stress in a random direction. (A) Schematic representations of the fitness landscape in a simplified 2D example similar to that in Figures 1D and 1E but projected to the plane defined by pathway activities. (B) Results of model simulations in a high-dimensional fitness space under conditions indicated in (A). Note that only the aneuploids selected by Stress X (the top 5% adaptive ones) are shown. A total of 1,000 cells were simulated. Each dot represents the relative fitness of an aneuploid cell compared to the euploid. Death zone was defined as having negative growth value.

(C) Six independent colonies isolated from radicicol-adapted population with gain of chrXV (as shown in Figure S3F) were grown under indicated conditions until saturation was reached in the fastest growing strain. Histograms show average amount of growth normalized to euploid and SEM derived from four replicates.

(D) chrlI and chrlX disomy strains generated previously by genetic manipulation (Torres et al., 2007) exhibit resistance to hygromycin B yet are sensitive to radicicol.

(E) Aneuploid strains generated by random triploid meiotic segregation with indicated karyotypic features were cultured in different concentrations of hygromycin B. Boxplots show growth relative to the euploid control with each dot representing an aneuploid strain. Karyotypes are categorized by their states of chrlI/X/XV dosage, but other chromosome aneuploidy may also be present in these strains. The amount of growth (OD increase) was normalized to the euploid with the nearest ploidy. The dashed line represents the average of normalized controls. \*\* indicates  $p < 0.01$  according to Mann-Whitney U test.

See also Figures S3 and S4.



**Figure 4. Different Sets of Genes on chrXV Cause Radicicol Resistance or Hygromycin B Sensitivity when Increased in Copy Number**

(A and B) Copy-number gain (+, by genomic integration) or loss (-, by genomic deletion) assays showing that increased copy numbers of *STI1* and *PDR5*, which are both critical for radicicol resistance, do not contribute to the hygromycin B hypersensitivity. Relative growth compared to the diploid control is reported in bar plots with the SEM derived from three replicates. Asterisks denote significant difference from the corresponding control (the diploid or chrXV trisomy [Tri-XV]) (\* $p < 0.05$ ; \*\* $p < 0.01$ ; two-tailed t test).

(legend continued on next page)

of a strategy to extinct such cell populations that accounts for their adaptive potential. Growth profiling of yeast aneuploids demonstrated that the phenotypic variation resulting from karyotype diversity scales directly with the degree of average inhibition of population growth under diverse stress conditions. Mathematical simulations based upon a few simple assumptions support the generality of our experimental observation. A more intuitive understanding of this phenomenon may lie in the highly pleiotropic nature of aneuploidy. The finding from yeast aneuploids was also recapitulated in human cancer cell lines derived from highly aneuploid solid tumors: the growth-suppressive effects of a majority of therapeutic compounds also significantly correlated with phenotypic heterogeneity in a positive manner. This correlation predicts that a heterogeneous aneuploid cancer population may harbor large phenotypic variation under stress, fueling rapid adaptive evolution. Frequent association between genetic heterogeneity and poor prognosis has indeed been observed in cancer clinics (Maley et al., 2006; McGranahan et al., 2012). Chromosomal instability (CIN) in mice also promotes tumor development (Sotillo et al., 2007; Weaver et al., 2007), although the extent to which genetic heterogeneity contributes to tumorigenesis in common mouse cancer models remains to be further explored. Beyond aneuploidy, the wide spectrum and large number of genetic modifications in a given cancer genome are likely to augment the pleiotropic effects on cellular pathways caused by aneuploidy.

Our experimental and theoretical findings argue on a general level that the difficulty of suppressing karyotypically heterogeneous cell populations is rooted in the large adaptive potential in the presence of severe stress. However, our findings do not rule out the possible existence of agents with broad inhibitory effects against aneuploids, especially when the karyotypic space is limited. The principle of ET in fact capitalizes on this notion by reducing karyotypic heterogeneity to an exceedingly narrow space through an evolutionary process. In our proof-of-principle experiments in budding yeast, this confined karyotypic space is simply gain of chrXV, which harbors the “driver mutations” for radicicol resistance (selection), whereas multiple “passenger” genes on chrXV confer enhanced sensitivity to hygromycin B, the agent for extinction, when increased in copy number (Figures 7C and 7D). The complex contribution of gene dosage to hygromycin B sensitivity and the requirement for multiple genes to achieve radicicol resistance predicted that an efficient escape from the ET is not easily attainable through single-gene mutations or copy-number changes. Although ET is not presumed to be an omnipotent solution for cancer or fungal infections, we argue that, by accounting for

the evolutionary trajectory of a cell population, ET may substantially reduce the risk of evolved drug resistance and disease relapse.

The principle of ET is distinct from the idea of targeting a single characteristic trait of aneuploids. We note that both drugs used in the yeast ET, radicicol and hygromycin B, can perturb proteome homeostasis (Orromendia et al., 2012; Singh et al., 1979; Taipale et al., 2010; Torres et al., 2007), but the responses of aneuploids to these drugs are karyotype specific rather than uniform. The rationale behind the dual drug treatments associated with ET anticipates evolutionary changes and is thus fundamentally different from the idea of synthetic lethality (Kaelin, 2005; Luo et al., 2009) or the use of dual drugs to target different aspects of a common defect of aneuploidy (Tang et al., 2011). ET drugs are not chosen to interfere with redundant processes but rather collaborate to force and intercept a predicted evolutionary trajectory of the genome. The same karyotype feature is selected by one drug but rejected by the other, as a result of linkage of two pools of genes, which respond oppositely to the two different drugs, on the same chromosome (Figure 7C).

Our experiments with *Candida* provide a possible ET example where PP, which selects against the gain of i(5L), could enhance the efficacy of fluconazole, the prior selection favoring i(5L) gain. PP is extremely well tolerated in the pediatric population for gastrointestinal treatment due to the near-zero absorption rate (Smith et al., 1976). Thus, PP may be used to augment fluconazole if the developed azole resistance is determined to be associated with i(5L) gain variants originated from gut flora.

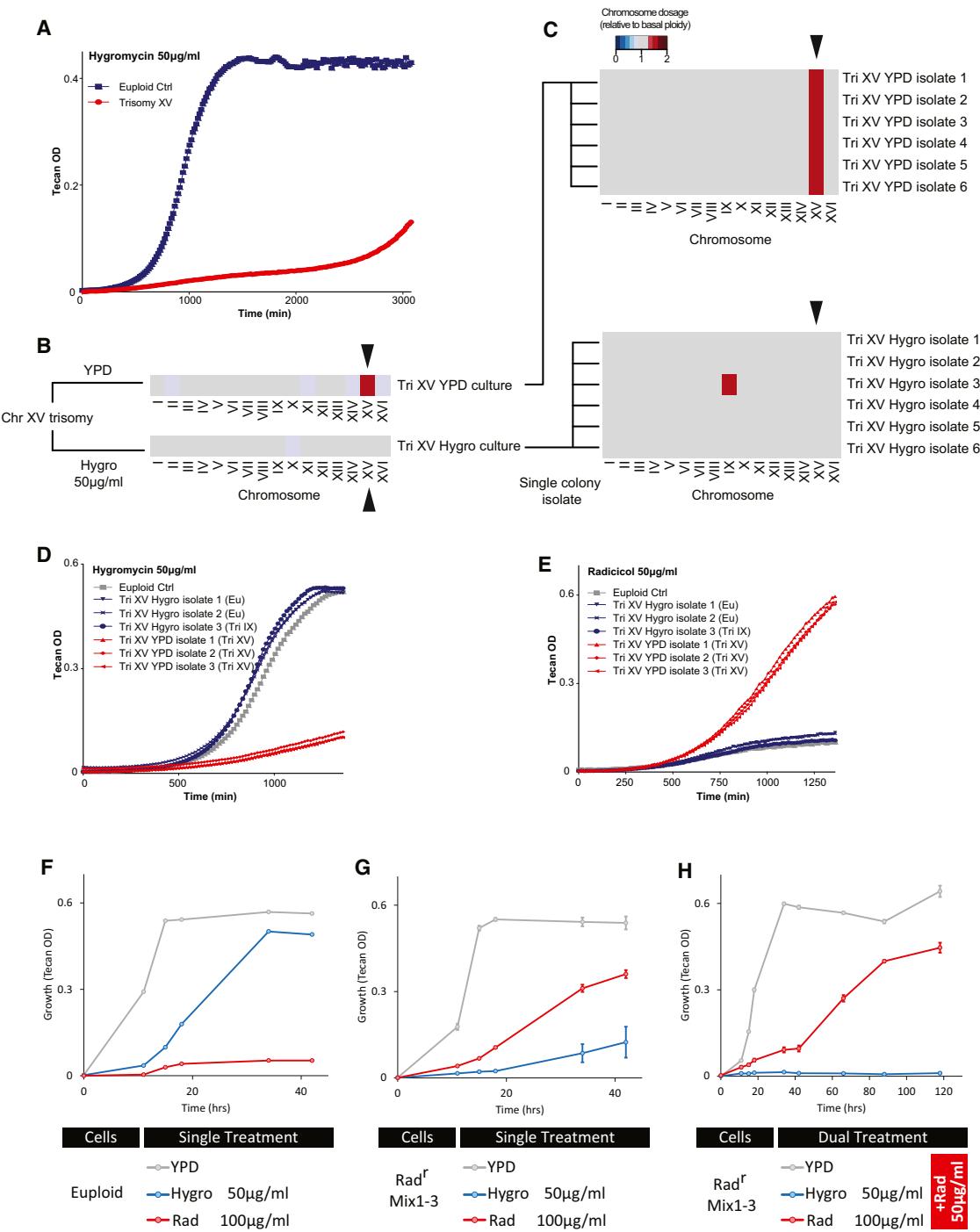
Our analysis of the cancer pharmacogenomics data suggests that irinotecan may form an ET with EGFR inhibitors against human cancers such as glioblastoma. Unlike other tumor types dependent on EGF signaling, glioblastoma responds poorly to EGFR inhibitors with the erlotinib response rate at <25% (Taylor et al., 2012). We envision that treatment with an EGFR inhibitor may further select for and thus increase the percentage of chr7-gained cells within the tumor mass. Supporting this idea, it was reported that EGF limitation selects for chr7 gain in human neural stem cells (Sareen et al., 2009). By analyzing existing pharmacogenomics data in brain tumor cell lines, we identified irinotecan, an FDA-approved chemical that targets topoisomerase 1, as a drug whose potency significantly correlates with chr7 gain. Additional experiments will be required to further confirm this correlation and test the efficacy of an ET formed with a combination of gefitinib or erlotinib with irinotecan against glioblastoma.

(C and D) Each of 453 genes located on chrXV was transformed into a diploid strain, and Z scores denoting the deviation of growth of each strain from the population average in the presence of 35 µg/ml hygromycin B were plotted against mRNA (C), using RNA-seq data, or protein expression abundance (D) (Ghaemmaghami et al., 2003) of each tested gene in the euploid S288c background. Note that protein abundance data were not retrieved for 30% genes (including CRS5 and RPS15). The gray area shows the 95% confidence interval for the linear fitting.

(E) Growth assays showing that copy-number increases (by genomic integration) of three genes (CRS5, RPS15, TRM11) on chrXV were individually sufficient in a diploid euploid context to reproduce enhanced sensitivity to hygromycin B, but not radicicol resistance, contrasting copy-number increase for ST11 and PDR5 as shown in (A).

(F) Growth assays showing that single-copy deletion of none of the three genes (CRS5, RPS15, TRM11) alone could rescue chrXV trisomy from hygromycin B hyper-sensitivity.

See also Figure S5.

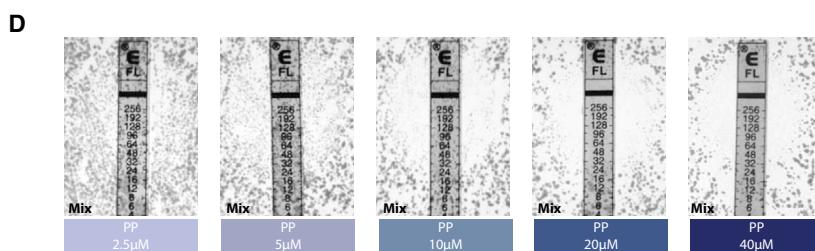
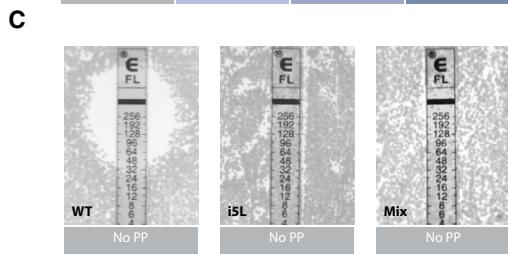
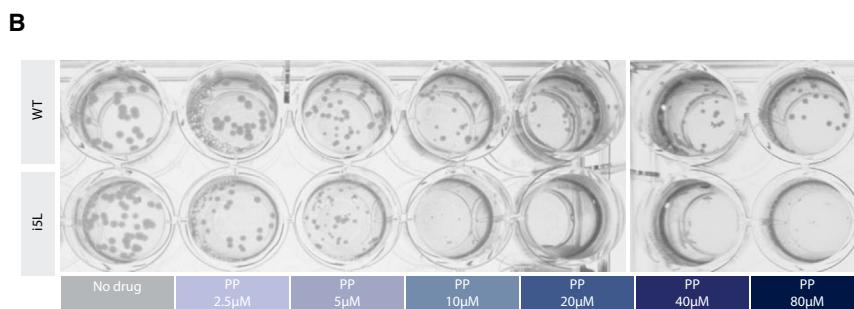
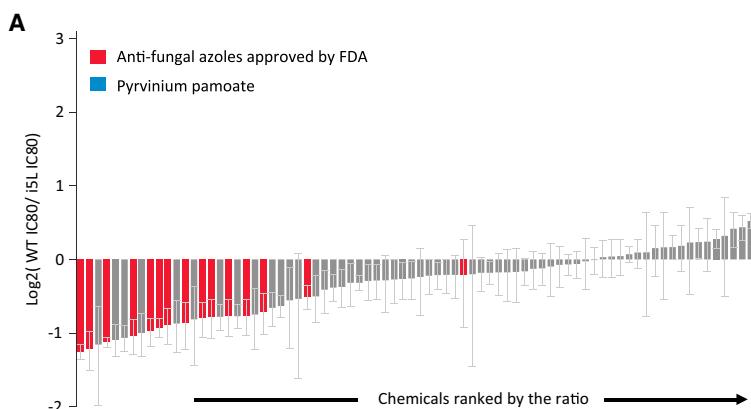


**Figure 5. The Combination of Radicicol and Hygromycin B Extincts Karyotypically Heterogeneous Cell Population**

(A–E) ChrXV trisomy was able to escape growth inhibition by hygromycin B through loss of the gained chrXV. (A) The growth (represented by OD reading on a Tecan reader) of both the euploid control and the trisomy XV strain was monitored in media containing 50  $\mu$ g/ml hygromycin B. (B) The additional copy of chrXV was lost in hygromycin B culture but not in YPD culture, as shown by the heatmap of karyotyping result of the final culture. (C) Karyotypes of six single colonies from the trisomy XV culture in YPD or hygromycin are shown, three of which were re-tested for growth in the presence of hygromycin B (D) or radicicol (E). Note that radicicol sensitivity was re-established in all three adapted colonies from the trisomy XV culture in hygromycin, whereas isolates from the YPD culture remained radicicol resistant.

(F–H) Combination of hygromycin B and radicicol effectively eradicates the radicicol-preselected aneuploid population. (F) Growth curves (as OD600 measured in Tecan) of the diploid control strain under conditions indicated. Note that 50  $\mu$ g/ml hygromycin B alone had milder growth suppression compared to 100  $\mu$ g/ml

(legend continued on next page)



## EXPERIMENTAL PROCEDURES

### Cell Culture

Budding yeast (strains are listed in Table S1, and plasmids/primers used to construct these strains are listed in Tables S2 and S3) was cultured using standard media. *C. albicans* cells recovered from frozen glycerol stocks were grown on YPD plates. For bulk liquid culture prior to drug-sensitivity assays, SC media with additional 80 mg/l uridine were used. RPMI1640 +

radicicol. (G) Growth curves of three populations pre-selected independently in the presence of radicicol (Figure S3E) under indicated conditions. (H) Growth curves of the same three populations as in (G) under indicated conditions where each drug was combined with 50 μg/ml radicicol. Each data point in (G) and (H) shows the mean and SEM from three experiments.

See also Figure S6.

## Figure 6. PP Effectively Targets the Fluconazole-Resistant *Candida* Aneuploid

(A) Relative IC<sub>80</sub> (80% inhibitory concentration) of the diploid versus the i5L Candida strain for each of the hits of the primary drug screen (Figure S7A). The error bars reflect the model-based standard error of fitting reported by drc package in R programming.  
 (B) Images of agar plates showing effectiveness of PP toward i5L candida.  
 (C) Resistance of the diploid, the i5L, or the i5L+diploid mix population toward fluconazole.  
 (D) PP at concentrations above 10 μM restored the sensitivity of the i5L+diploid mix population toward fluconazole in the E-test, in accordance to its singular form's activity against the i5L strain shown in (B). Note that even though the initial plating density was the same, due to the inhibition of the i5L cells, the overall growth was less in (D) compared to (C). Note that our euploid strain also exhibited a reduced susceptibility to fluconazole compared to the clinical E-test standard strain, which may be attributed to other point mutations (such as the hyperactive TAC1) within this strain (Selmecki et al., 2006). All plate images were taken after 48 hr culture.  
 See also Figure S7.

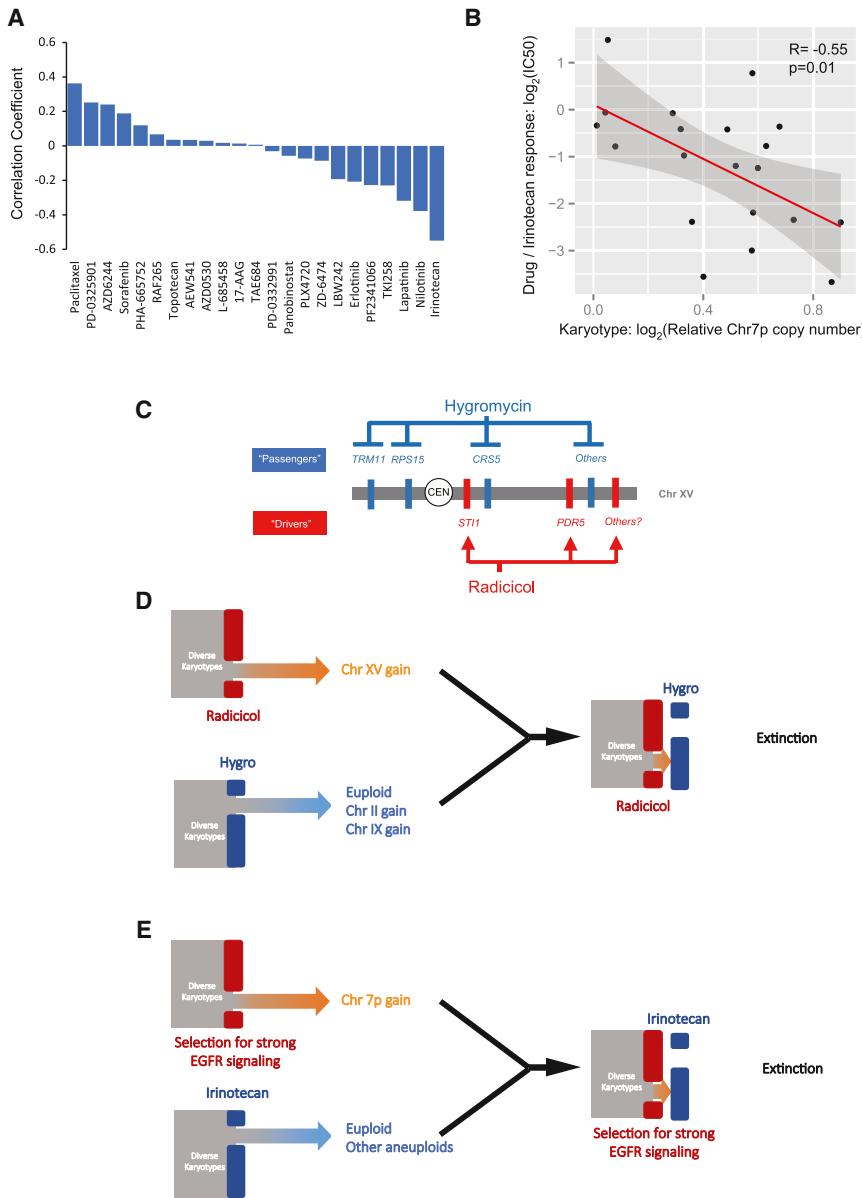
0.165M MOPS buffer + 0.2% glucose without bicarbonate (Lonza) was used for 96-well culture according to the Clinical Laboratory and Standards Institute (CLSI) standard M27-A3. For 384-well cultures, an additional 2% glucose was added to accelerate the growth. Agar plates were prepared using RPMI1640 with MOPS without bicarbonate powder (US Biological) supplemented with 2% glucose and 1.5% agarose, following the E-test strip manufacturer's guideline (bioMerieux).

### Yeast Growth Assays

For continuous OD monitoring (e.g., Figure 3D), cultures were set up in 96-well plates sealed with parafilm in a Tecan M200Pro reader with orbital shaking. OD at wavelength 595 nm was taken every 15 min and analyzed by using the Magellan 7 software (Tecan). For cultures whose OD reading was taken intermittently every several hours (e.g., Figure 5F), the culture plates were placed in a non-shaking incubator within a humid chamber before each reading after agitation. The OD at wavelength 595 nm

was recorded by using a Tecan M200Pro reader, and the data files were processed in R. The growth assays lasted until the fastest growing culture reached saturation, at which time the last OD readings of all strains were recorded.

The drug concentrations of hygromycin B and radicicol were adjusted for different media (e.g., YPD versus SC-ura) and/or the strain background, so that the wild-type control showed the same growth delay as was observed for the strain RLY2628 in YPD media containing the stated drug concentration.



**Figure 7. A Potential Drug Targeting chr7 Gain in Brain Cancer and Schematic Summary of the Mechanism and Principle of ET using the Yeast Example**

(A) Correlation coefficients of drug response ( $\text{IC}_{50}$ ) with chr7p dosage in 29 CNS tumor cell lines across 23 different therapeutic compounds were plotted as bar graphs.

(B) The dot plot illustrates the details of the correlation between dosage of chr7p and sensitivity to irinotecan, with each dot showing the drug response and chr7p dosage of each cell line. The red line shows linear fitting, and the gray area shows the fitting range with 95% confidence interval. Note that a total of 20 cell lines were included here, as the  $\text{IC}_{50}$  data for nine cell lines were not available for irinotecan.

(C) The molecular makeup of the ET against aneuploid yeast.

(D) Schematic summary of opposing selective effects of radicicol and hygromycin B on chrXV gain impose an adaptive dilemma for the yeast heterogeneous aneuploid population.

(E) An ET may be established against glioblastoma by opposing selective effects on chr7p gain.

The growth of genetic variants was normalized to the corresponding wild-type controls.

A comprehensive description of all methods used can be found in [Extended Experimental Procedures](#).

#### ACCESSION NUMBERS

All screening data (on both budding yeast and *C. albicans*) will be deposited into Open Data Repository (ODR) of the Stowers Institute for Medical Research and become openly available (<http://odr.stowers.org/websimr/>).

#### SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, three tables, and seven figures and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2015.01.026>.

#### AUTHOR CONTRIBUTIONS

R.L. and G.C. designed the project. G.C. and W.A.M. performed most budding yeast experiments. G.C. and H.C. performed most *Candida* experiments with help from J.B. A.K., B.R., G.C., and R.L. built the model. G.C. and A.K. performed analyses of the human cancer cell line data. J.C., S.M. and L.W. carried out parts of the budding yeast and *Candida* screens. W.D.B. performed qPCR karyotyping. J.S.H. assisted with the flow cytometry. C.W.S. assisted with the RNA-seq data analysis. G.C. and R.L. prepared the manuscript. R.L. supervised the project.

#### ACKNOWLEDGMENTS

We thank A. Amon for providing the set of disomy aneuploid yeast strains and J. Gerton for sharing haploid yeast RNA-seq data. This work was supported by NIH grant RO1GM059964 to R.L.

Received: July 31, 2014  
 Revised: November 17, 2014  
 Accepted: December 31, 2014  
 Published: February 12, 2015

## REFERENCES

Barretina, J., Caponigro, G., Stransky, N., Venkatesan, K., Margolin, A.A., Kim, S., Wilson, C.J., Lehár, J., Kryukov, G.V., Sonkin, D., et al. (2012). The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* 483, 603–607.

Bean, J., Brennan, C., Shih, J.-Y., Riely, G., Viale, A., Wang, L., Chitale, D., Motoi, N., Szoke, J., Broderick, S., et al. (2007). MET amplification occurs with or without T790M mutations in EGFR mutant lung tumors with acquired resistance to gefitinib or erlotinib. *Proc. Natl. Acad. Sci. USA* 104, 20932–20937.

Beroukhim, R., Mermel, C.H., Porter, D., Wei, G., Raychaudhuri, S., Donovan, J., Barretina, J., Boehm, J.S., Dobson, J., Urashima, M., et al. (2010). The landscape of somatic copy-number alteration across human cancers. *Nature* 463, 899–905.

Borneman, A.R., Desany, B.A., Riches, D., Affourtit, J.P., Forgan, A.H., Pretorius, I.S., Egholm, M., and Chambers, P.J. (2011). Whole-genome comparison reveals novel genetic elements that characterize the genome of industrial strains of *Saccharomyces cerevisiae*. *PLoS Genet.* 7, e1001287.

Borovinskaya, M.A., Shoji, S., Fredrick, K., and Cate, J.H.D. (2008). Structural basis for hygromycin B inhibition of protein biosynthesis. *RNA* 14, 1590–1599.

Bredel, M., Scholtens, D.M., Harsh, G.R., Bredel, C., Chandler, J.P., Renfrow, J.J., Yadav, A.K., Vogel, H., Scheck, A.C., Tibshirani, R., and Sikic, B.I. (2009). A network model of a cooperative genetic landscape in brain tumors. *JAMA* 302, 261–275.

Burrell, R.A., McGranahan, N., Bartek, J., and Swanton, C. (2013). The causes and consequences of genetic heterogeneity in cancer evolution. *Nature* 501, 338–345.

St Charles, J., Hamilton, M.L., and Petes, T.D. (2010). Meiotic chromosome segregation in triploid strains of *Saccharomyces cerevisiae*. *Genetics* 186, 537–550.

Chen, G., Bradford, W.D., Seidel, C.W., and Li, R. (2012a). Hsp90 stress potentiates rapid cellular adaptation through induction of aneuploidy. *Nature* 482, 246–250.

Chen, G., Rubinstein, B., and Li, R. (2012b). Whole chromosome aneuploidy: big mutations drive adaptation by phenotypic leap. *Bioessays* 34, 893–900.

Davoli, T., Xu, A.W., Mengwasser, K.E., Sack, L.M., Yoon, J.C., Park, P.J., and Elledge, S.J. (2013). Cumulative haploinsufficiency and triplosensitivity drive aneuploidy patterns and shape the cancer genome. *Cell* 155, 948–962.

Gerlinger, M., Rowan, A.J., Horswell, S., Larkin, J., Endesfelder, D., Gronroos, E., Martinez, P., Matthews, N., Stewart, A., Tarpey, P., et al. (2012). Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N. Engl. J. Med.* 366, 883–892.

Ghaemmaghami, S., Huh, W.-K., Bower, K., Howson, R.W., Belle, A., Dephoure, N., O'Shea, E.K., and Weissman, J.S. (2003). Global analysis of protein expression in yeast. *Nature* 425, 737–741.

Harrison, B.D., Hashemi, J., Bibi, M., Pulver, R., Bavli, D., Nahmias, Y., Wellington, M., Sapiro, G., and Berman, J. (2014). A tetraploid intermediate precedes aneuploid formation in yeasts exposed to fluconazole. *PLoS Biol.* 12, e1001815.

Heiser, L.M., Sadanandam, A., Kuo, W.-L., Benz, S.C., Goldstein, T.C., Ng, S., Gibb, W.J., Wang, N.J., Ziyad, S., Tong, F., et al. (2012). Subtype and pathway specific responses to anticancer compounds in breast cancer. *Proc. Natl. Acad. Sci. USA* 109, 2724–2729.

Ho, C.H., Magtanong, L., Barker, S.L., Gresham, D., Nishimura, S., Natarajan, P., Koh, J.L.Y., Porter, J., Gray, C.A., Andersen, R.J., et al. (2009). A molecular barcoded yeast ORF library enables mode-of-action analysis of bioactive compounds. *Nat. Biotechnol.* 27, 369–377.

Holland, A.J., and Cleveland, D.W. (2009). Boveri revisited: chromosomal instability, aneuploidy and tumorigenesis. *Nat. Rev. Mol. Cell Biol.* 10, 478–487.

Hughes, T.R., Roberts, C.J., Dai, H., Jones, A.R., Meyer, M.R., Slade, D., Burchard, J., Dow, S., Ward, T.R., Kidd, M.J., et al. (2000). Widespread aneuploidy revealed by DNA microarray expression profiling. *Nat. Genet.* 25, 333–337.

Infante, J.J., Dombek, K.M., Rebordinos, L., Cantoral, J.M., and Young, E.T. (2003). Genome-wide amplifications caused by chromosomal rearrangements play a major role in the adaptive evolution of natural yeast. *Genetics* 165, 1745–1759.

Jones, L., Wei, G., Sevcikova, S., Phan, V., Jain, S., Shieh, A., Wong, J.C.Y., Li, M., Dubansky, J., Maunakea, M.L., et al. (2010). Gain of MYC underlies recurrent trisomy of the MYC chromosome in acute promyelocytic leukemia. *J. Exp. Med.* 207, 2581–2594.

Kaelin, W.G., Jr. (2005). The concept of synthetic lethality in the context of anti-cancer therapy. *Nat. Rev. Cancer* 5, 689–698.

Kvitek, D.J., Will, J.L., and Gasch, A.P. (2008). Variations in stress sensitivity and genomic expression in diverse *S. cerevisiae* isolates. *PLoS Genet.* 4, e1000223.

Lee, A.J.X., Endesfelder, D., Rowan, A.J., Walther, A., Birkbak, N.J., Futreal, P.A., Downward, J., Szallasi, Z., Tomlinson, I.P.M., Howell, M., et al. (2011). Chromosomal instability confers intrinsic multidrug resistance. *Cancer Res.* 71, 1858–1870.

Leprohon, P., Légaré, D., Raymond, F., Madore, É., Hardiman, G., Corbeil, J., and Ouellette, M. (2009). Gene expression modulation is associated with gene amplification, supernumerary chromosomes and chromosome loss in antimony-resistant *Leishmania infantum*. *Nucleic Acids Res.* 37, 1387–1399.

Lewontin, R.C. (1970). The units of selection. *Annu. Rev. Ecol. Syst.* 1, 1–18.

Llewellyn, M.S., Rivett-Carnac, J.B., Fitzpatrick, S., Lewis, M.D., Yeo, M., Gaunt, M.W., and Miles, M.A. (2011). Extraordinary *Trypanosoma cruzi* diversity within single mammalian reservoir hosts implies a mechanism of diversifying selection. *Int. J. Parasitol.* 41, 609–614.

Luo, J., Emanuele, M.J., Li, D., Creighton, C.J., Schlabaach, M.R., Westbrook, T.F., Wong, K.-K., and Elledge, S.J. (2009). A genome-wide RNAi screen identifies multiple synthetic lethal interactions with the Ras oncogene. *Cell* 137, 835–848.

Maley, C.C., Galipeau, P.C., Finley, J.C., Wongsurawat, V.J., Li, X., Sanchez, C.A., Paulson, T.G., Blount, P.L., Risques, R.A., Rabinovitch, P.S., and Reid, B.J. (2006). Genetic clonal diversity predicts progression to esophageal adenocarcinoma. *Nat. Genet.* 38, 468–473.

Mannaert, A., Downing, T., Imamura, H., and Dujardin, J.-C. (2012). Adaptive mechanisms in pathogens: universal aneuploidy in *Leishmania*. *Trends Parasitol.* 28, 370–376.

Marichal, P., Vanden Bossche, H., Odds, F.C., Nobels, G., Warnock, D.W., Timmerman, V., Van Broeckhoven, C., Fay, S., and Mose-Larsen, P. (1997). Molecular biological characterization of an azole-resistant *Candida glabrata* isolate. *Antimicrob. Agents Chemother.* 41, 2229–2237.

Marr, K.A., White, T.C., van Burik, J.-A.H., and Bowden, R.A. (1997). Development of fluconazole resistance in *Candida albicans* causing disseminated infection in a patient undergoing marrow transplantation. *Clin. Infect. Dis.* 25, 908–910.

McGranahan, N., Burrell, R.A., Endesfelder, D., Novelli, M.R., and Swanton, C. (2012). Cancer chromosomal instability: therapeutic and diagnostic challenges. *EMBO Rep.* 13, 528–538.

Merlo, L.M.F., Pepper, J.W., Reid, B.J., and Maley, C.C. (2006). Cancer as an evolutionary and ecological process. *Nat. Rev. Cancer* 6, 924–935.

Minning, T.A., Weatherly, D.B., Flibotte, S., and Tarleton, R.L. (2011). Widespread, focal copy number variations (CNV) and whole chromosome aneuploidies in *Trypanosoma cruzi* strains revealed by array comparative genomic hybridization. *BMC Genomics* 12, 139.

Mitelman, F., Johansson, B., and Mertens, F. (eds.) (2012). Mitelman Database of Chromosome Aberrations and Gene Fusions in Cancer. <http://cgap.nci.nih.gov/Chromosomes/Mitelman>.

Navin, N., Kendall, J., Troge, J., Andrews, P., Rodgers, L., McIndoo, J., Cook, K., Stepansky, A., Levy, D., Esposito, D., et al. (2011). Tumour evolution inferred by single-cell sequencing. *Nature* 472, 90–94.

Network, T.C.G.A.R.; Cancer Genome Atlas Research Network (2008). Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* 455, 1061–1068.

Ng, A.P., Hyland, C.D., Metcalf, D., Carmichael, C.L., Loughran, S.J., Di Rago, L., Kile, B.T., and Alexander, W.S. (2010). Trisomy of Erg is required for myeloproliferation in a mouse model of Down syndrome. *Blood* 115, 3966–3969.

Ni, M., Feretzaki, M., Li, W., Floyd-Averette, A., Mieczkowski, P., Dietrich, F.S., and Heitman, J. (2013). Unisexual and heterosexual meiotic reproduction generate aneuploidy and phenotypic diversity de novo in the yeast *Cryptococcus neoformans*. *PLoS Biol.* 11, e1001653.

Okoli, I., Coleman, J.J., Tempakakis, E., An, W.F., Holson, E., Wagner, F., Conery, A.L., Larkins-Ford, J., Wu, G., Stern, A., et al. (2009). Identification of antifungal compounds active against *Candida albicans* using an improved high-throughput *Caenorhabditis elegans* assay. *PLoS ONE* 4, e7025.

Oromendia, A.B., and Amon, A. (2014). Aneuploidy: implications for protein homeostasis and disease. *Dis. Model. Mech.* 7, 15–20.

Oromendia, A.B., Dodgson, S.E., and Amon, A. (2012). Aneuploidy causes proteotoxic stress in yeast. *Genes Dev.* 26, 2696–2708.

Pavelka, N., Rancati, G., Zhu, J., Bradford, W.D., Saraf, A., Florens, L., Sanderson, B.W., Hattem, G.L., and Li, R. (2010). Aneuploidy confers quantitative proteome changes and phenotypic variation in budding yeast. *Nature* 468, 321–325.

Rancati, G., Pavelka, N., Fleharty, B., Noll, A., Trimble, R., Walton, K., Perera, A., Staehling-Hampton, K., Seidel, C.W., and Li, R. (2008). Aneuploidy underlies rapid adaptive evolution of yeast cells deprived of a conserved cytokinesis motor. *Cell* 135, 879–893.

Sareen, D., McMillan, E., Ebert, A.D., Shelley, B.C., Johnson, J.A., Meisner, L.F., and Svendsen, C.N. (2009). Chromosome 7 and 19 trisomy in cultured human neural progenitor cells. *PLoS ONE* 4, e7630.

Selmecki, A., Forche, A., and Berman, J. (2006). Aneuploidy and isochromosome formation in drug-resistant *Candida albicans*. *Science* 313, 367–370.

Selmecki, A., Gerami-Nejad, M., Paulson, C., Forche, A., and Berman, J. (2008). An isochromosome confers drug resistance in vivo by amplification of two genes, ERG11 and TAC1. *Mol. Microbiol.* 68, 624–641.

Selmecki, A.M., Dulmage, K., Cowen, L.E., Anderson, J.B., and Berman, J. (2009). Acquisition of aneuploidy provides increased fitness during the evolution of antifungal drug resistance. *PLoS Genet.* 5, e1000705.

Singh, A., Ursic, D., and Davies, J. (1979). Phenotypic suppression and misreading *Saccharomyces cerevisiae*. *Nature* 277, 146–148.

Sionov, E., Lee, H., Chang, Y.C., and Kwon-Chung, K.J. (2010). *Cryptococcus neoformans* overcomes stress of azole drugs by formation of disomy in specific multiple chromosomes. *PLoS Pathog.* 6, e1000848.

Smith, T.C., Kinkel, A.W., Gryczko, C.M., and Goulet, J.R. (1976). Absorption of pyrvinium pamoate. *Clin. Pharmacol. Ther.* 19, 802–806.

Sotillo, R., Hernando, E., Díaz-Rodríguez, E., Teruya-Feldstein, J., Cordón-Cardo, C., Lowe, S.W., and Benezra, R. (2007). Mad2 overexpression promotes aneuploidy and tumorigenesis in mice. *Cancer Cell* 11, 9–23.

Sotillo, R., Schwartzman, J.-M., Soccia, N.D., and Benezra, R. (2010). Mad2-induced chromosome instability leads to lung tumour relapse after oncogene withdrawal. *Nature* 464, 436–440.

Sottoriva, A., Spiteri, I., Piccirillo, S.G.M., Touloumis, A., Collins, V.P., Marioni, J.C., Curtis, C., Watts, C., and Tavaré, S. (2013). Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. *Proc. Natl. Acad. Sci. USA* 110, 4009–4014.

Spitzer, M., Griffiths, E., Blakely, K.M., Wildenhain, J., Ejim, L., Rossi, L., De Pascale, G., Curak, J., Brown, E., Tyers, M., and Wright, G.D. (2011). Cross-species discovery of synthetic drug combinations that potentiate the anti-fungal fluconazole. *Mol. Syst. Biol.* 7, 499.

Taipale, M., Jarosz, D.F., and Lindquist, S. (2010). HSP90 at the hub of protein homeostasis: emerging mechanistic insights. *Nat. Rev. Mol. Cell Biol.* 11, 515–528.

Tang, Y.-C., Williams, B.R., Siegel, J.J., and Amon, A. (2011). Identification of aneuploidy-selective antiproliferation compounds. *Cell* 144, 499–512.

Taylor, T.E., Furnari, F.B., and Cavenee, W.K. (2012). Targeting EGFR for treatment of glioblastoma: molecular basis to overcome resistance. *Curr. Cancer Drug Targets* 12, 197–209.

Torres, E.M., Sokolsky, T., Tucker, C.M., Chan, L.Y., Boselli, M., Dunham, M.J., and Amon, A. (2007). Effects of aneuploidy on cellular physiology and cell division in haploid yeast. *Science* 317, 916–924.

Ubeda, J.-M., Légaré, D., Raymond, F., Ouameur, A.A., Boisvert, S., Rigault, P., Corbeil, J., Tremblay, M.J., Olivier, M., Papadopoulou, B., and Ouellette, M. (2008). Modulation of gene expression in drug resistant *Leishmania* is associated with gene amplification, gene deletion and chromosome aneuploidy. *Genome Biol.* 9, R115.

Wang, Y., Waters, J., Leung, M.L., Unruh, A., Roh, W., Shi, X., Chen, K., Scheet, P., Vattathil, S., Liang, H., et al. (2014). Clonal evolution in breast cancer revealed by single nucleus genome sequencing. *Nature* 512, 155–160.

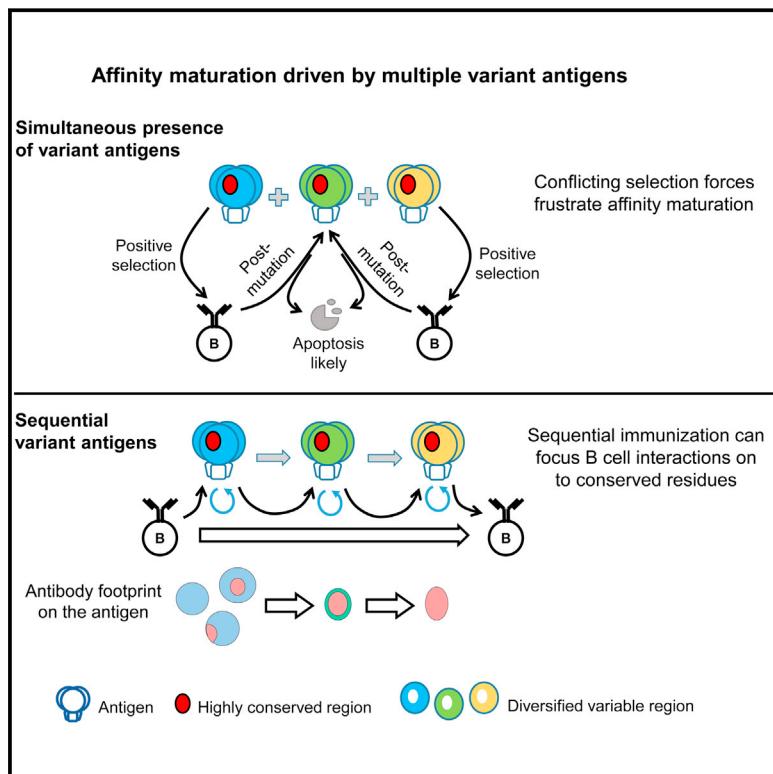
Weaver, B.A.A., Silk, A.D., Montagna, C., Verdier-Pinard, P., and Cleveland, D.W. (2007). Aneuploidy acts both oncogenically and as a tumor suppressor. *Cancer Cell* 11, 25–36.

Wheeler, D.L., Dunn, E.F., and Harari, P.M. (2010). Understanding resistance to EGFR inhibitors—impact on future treatment strategies. *Nature reviews. Clin. Oncol.* 7, 493–507.

Zhu, J., Pavelka, N., Bradford, W.D., Rancati, G., and Li, R. (2012). Karyotypic determinants of chromosome instability in aneuploid budding yeast. *PLoS Genet.* 8, e1002719.

# Manipulating the Selection Forces during Affinity Maturation to Generate Cross-Reactive HIV Antibodies

## Graphical Abstract



## Authors

Shenshen Wang, Jordi Mata-Fink, ..., Mehran Kardar, Arup K. Chakraborty

## Correspondence

arupc@mit.edu

## In Brief

In silico model of antibody affinity maturation explains why antibodies capable of cross-reacting with different variants of viral antigens are uncommon and shows that sequential immunization is effective at inducing cross-reactive HIV antibodies focused on the shared CD4 binding site.

## Highlights

- In silico model of affinity maturation driven by variant antigens
- Conflicting selection forces due to antigen variants can frustrate maturation
- Key variables that control evolution of cross-reactive antibodies identified
- Sequential immunization favored for inducing cross-reactive antibodies

# Manipulating the Selection Forces during Affinity Maturation to Generate Cross-Reactive HIV Antibodies

Shenshen Wang,<sup>1,2,3</sup> Jordi Mata-Fink,<sup>1,2</sup> Barry Kriegsman,<sup>2</sup> Melissa Hanson,<sup>5</sup> Darrell J. Irvine,<sup>1,5,8</sup> Herman N. Eisen,<sup>8,9</sup> Dennis R. Burton,<sup>1,4,10</sup> K. Dane Wittrup,<sup>2,5,10</sup> Mehran Kardar,<sup>3,10</sup> and Arup K. Chakraborty<sup>1,2,3,5,6,7,10,\*</sup>

<sup>1</sup>Ragon Institute of MGH, MIT, and Harvard, Cambridge, MA 02139

<sup>2</sup>Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139

<sup>3</sup>Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139

<sup>4</sup>Department of Immunology and Microbial Science, The Scripps Research Institute, La Jolla, CA 92037

<sup>5</sup>Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139

<sup>6</sup>Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139

<sup>7</sup>Institute for Medical Engineering and Science, Massachusetts Institute of Technology, Cambridge, MA 02139

<sup>8</sup>Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA 02139

<sup>9</sup>Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139

<sup>10</sup>Co-senior author

\*Correspondence: arupc@mit.edu

<http://dx.doi.org/10.1016/j.cell.2015.01.027>

## SUMMARY

Generation of potent antibodies by a mutation-selection process called affinity maturation is a key component of effective immune responses. Antibodies that protect against highly mutable pathogens must neutralize diverse strains. Developing effective immunization strategies to drive their evolution requires understanding how affinity maturation happens in an environment where variants of the same antigen are present. We present an in silico model of affinity maturation driven by antigen variants which reveals that induction of cross-reactive antibodies often occurs with low probability because conflicting selection forces, imposed by different antigen variants, can frustrate affinity maturation. We describe how variables such as temporal pattern of antigen administration influence the outcome of this frustrated evolutionary process. Our calculations predict, and experiments in mice with variant gp120 constructs of the HIV envelope protein confirm, that sequential immunization with antigen variants is preferred over a cocktail for induction of cross-reactive antibodies focused on the shared CD4 binding site epitope.

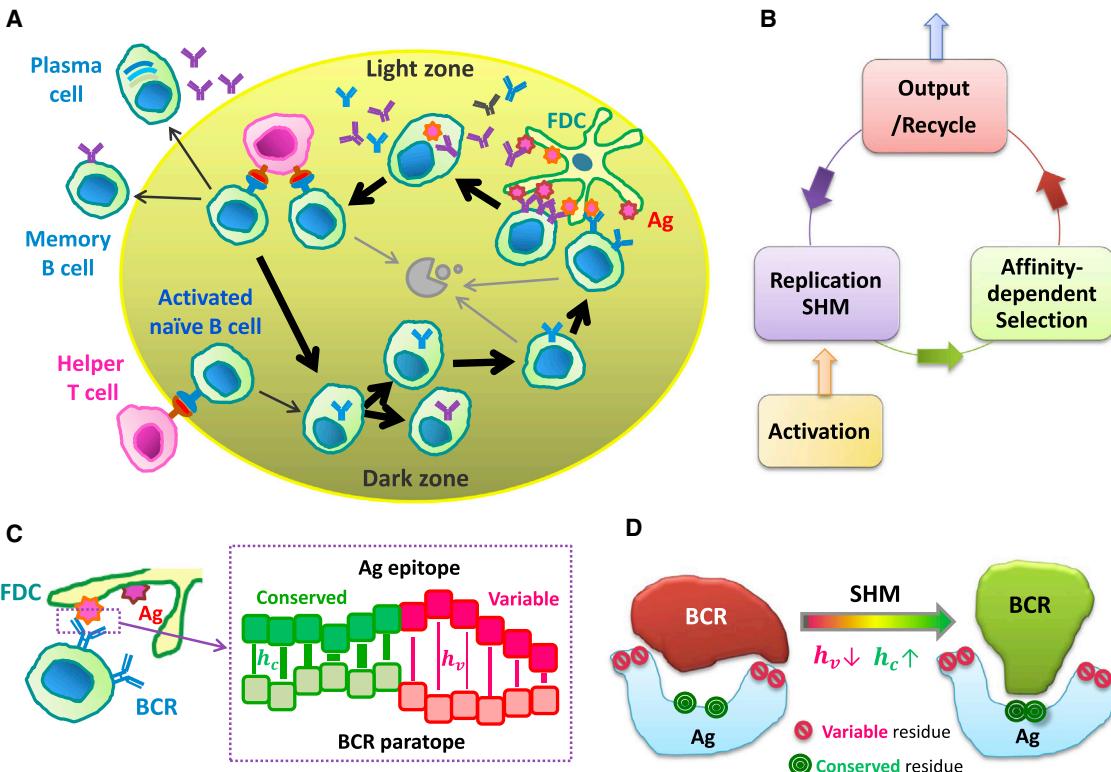
## INTRODUCTION

Antibodies (Abs) with high affinity for antigen are produced by the process of affinity maturation (AM), which takes place in germinal centers (GCs). GCs are dynamic structures within secondary lymphoid tissues that arise in response to antigen stimulation (Shlomchik and Weisel, 2012; Victora and Nussenzweig, 2012).

GCs house B cells, antigen-specific T helper cells that develop in concert with GC B cells (Baumjohann et al., 2013; Kelsoe, 1996) and antigens presented on follicular dendritic cells (FDCs) (Figure 1A). GC B cells enhance the antigen affinity of their receptors by 10- to 1,000-fold through cycles of mutation and selection against antigens presented on FDCs, a Darwinian evolutionary process that occurs on a very short timescale. Soluble forms of the high-affinity receptors are potent Abs. AM has been studied extensively using diverse experimental methods (Batista and Neuberger, 1998; Berek and Milstein, 1987; Berek et al., 1991; Eisen and Siskind, 1964; Jacob et al., 1991; Kocks and Rajewsky, 1988), mathematical models (Deem and Lee, 2003; Kepler and Perelson, 1993; Meyer-Hermann, 2002; Meyer-Hermann et al., 2006; Oprea and Perelson, 1997; Zhang and Shakhnovich, 2010), and computer simulations (Keşmir and De Boer, 2003; Shlomchik et al., 1998; Swerdlin et al., 2008). Recent experiments have uncovered new aspects of GC dynamics (Allen et al., 2007; Shulman et al., 2013; Victora et al., 2010).

Effective Ab responses are likely to be required for a protective prophylactic vaccine against highly mutable pathogens. For HIV, a quintessential example of such a pathogen, there has been no definitive success in designing such a vaccine, despite extensive efforts (Burton et al., 2012; Klein et al., 2013; Kwong et al., 2013; Mascola and Haynes, 2013). This is in large part due to the variability of HIV (Burton et al., 2012; Mascola and Haynes, 2013).

The protective effects of an Ab are predicated on its ability to bind to a set of residues (the epitope) on the surface of virions. For example, the HIV-1 envelope glycoprotein trimer (Env) is the sole target of known HIV-1 neutralizing Abs (Burton et al., 2012; Kwong et al., 2013; Mascola and Haynes, 2013; West et al., 2014). During the first few months of infection, the induced neutralizing Abs are primarily strain-specific and target variable regions of Env. As the concentrations of such strain specific Abs increase, neutralization escape variants with mutations in the variable epitopes are selected. A pathogen's molecular



**Figure 1. Schematic Depiction of In Silico Model**

(A) Players and processes in the GCR.  
 (B) Major steps in our in silico model of the GCR.  
 (C) Model for BCR-Ag interactions. Left: A BCR interacting with an FDC-held Ag. Right: A zoom-in view of interactions (bars) between the residues on the BCR paratope and those on the Ag epitope. An affinity-affecting mutation on a paratope residue will change its interaction strength  $h$  with the corresponding epitope residue, denoted by  $h_c$  ( $h_v$ ) if the latter is conserved (variable).  
 (D) Conformational coupling between residues on the BCR is incorporated via correlated changes in  $h_c$  and  $h_v$ ; weakening interaction with the variable residues and the residues that shield the conserved residues of the epitope (red symbols), i.e.,  $h_v \downarrow$ , would facilitate access to the conserved residues (green symbols), i.e.,  $h_c \uparrow$ .  
 See also Figure S1.

surfaces that interact with host receptors (e.g., the CD4 binding site) provide epitopes that contain a relatively conserved set of residues. They can serve as targets for Abs that are able to neutralize a greater diversity of HIV strains. A shield of glycans and immunodominant variable loops can restrict antibody access to these relatively conserved epitopes (Julien et al., 2013; Kwong et al., 2002; Lyumkis et al., 2013; Pancera et al., 2014; Wei et al., 2003; Wyatt et al., 1998). Yet, some HIV-1 infected individuals do develop Abs that focus on such epitopes and neutralize a broad cross section of HIV strains *in vitro* (Kwong et al., 2013). However, these broadly neutralizing antibodies (bnAbs) are produced only beginning around 2 years after infection and in only a limited fraction of chronically infected patients. Nonetheless, the emergence of bnAbs in some patients is proof that AM can result in bnAbs. This suggests the tantalizing possibility that appropriately designed immunogens and immunization protocols may be able to elicit bnAbs rapidly in a large fraction of individuals, thus resulting in an effective vaccine.

Recent longitudinal tracking of a developing HIV-1 bnAb lineage and the co-evolving virus in a patient showed that extensive,

and even specifically directed, viral diversification occurred prior to the development of breadth (cross-reactive Abs) (Gao et al., 2014; Liao et al., 2013), supporting the idea that escape mutants of the virus drive bnAb evolution. More importantly, this observation highlights that induction of bnAbs will likely require immunization with multiple variants of the antigen.

When multiple complex antigen variants are used as immunogens, several new questions become important, for example: (1) which antigen variants should be used as immunogens; (2) what should be the concentrations and temporal order in which they are administered (e.g., cocktails versus sequential)? The answers to these questions are drawn from a huge number of possibilities, and random selection from a large number of combinations of options may not allow sufficient sampling to find efficacious strategies. Intuition is unlikely to guide choices correctly because a mechanistic understanding of how AM occurs in the face of variant or mutating antigens is not available, as past studies have focused on AM in response to single model antigens. This gap in basic immunobiology needs to be addressed. A fundamental understanding of AM induced by several antigen

variants that mimic the complexity of how conserved epitopes are shielded on the intact viral spike could be harnessed to design optimal immunogens and immunization protocols for the development of universal vaccines against highly mutable pathogens (e.g., HIV-1, influenza).

To take steps toward these goals, we developed an *in silico* model of AM induced by variant antigens. This stochastic dynamic model enables examination of the key mechanisms and factors that influence AM in the face of variant antigens and the development of bnAbs. Our calculations predict markedly distinct outcomes if designed antigen (Ag) variants are presented in different concentrations and temporal patterns during immunization. Experiments in mice using model HIV antigens are consistent with these predictions. Thus, our complementary computational and experimental results provide broadly applicable fundamental insights, and a guide for further studies aimed at overcoming roadblocks to the induction of bnAbs against HIV-1 by vaccination.

## RESULTS

### In Silico Model

The purpose of our *in silico* modeling is not quantitative recapitulation of existing experiments but to provide fundamental mechanistic insights into AM induced by multiple Ag variants and to compare the predicted relative efficacy of different immunization schemes in inducing cross-reactive broadly neutralizing Abs (bnAbs). Two key features not considered before must be incorporated explicitly in such studies; viz., the presence of multiple Ag variants during GC reactions and the molecular complexity of the Ags. To mimic interactions of B cell receptors with complex immunogens such as the HIV-1 trimeric spike, we account for the fact that the conserved protein epitope of desired Abs are partially shielded, and that insertions and deletions in the variable loops can hinder the formation of strong interactions with the conserved residues of the epitopes (e.g., those in the CD4bs).

We computationally simulate the dynamics of a typical GC (Figure 1A). The GC reaction (GCR) starts on day 3 (Nieuwenhuis and Opstelten, 1984) after Ag injection, with three B cell blasts (Jacob et al., 1991; Kroese et al., 1987) that barely meet a relatively low-affinity threshold,  $E_a$ . The qualitative results we report do not depend on the particular value of  $E_a$  since all other affinities are scaled relative to it. The B cells expand without mutation, reaching a population size of about 1,500 cells around day 7 (Jacob et al., 1991).

### Hypermutation in the Dark Zone of the GC

At day 7, AID-mediated somatic hypermutation (SHM) in the immunoglobulin (Ig) genes turns on (Källberg et al., 1996) with a rate of  $10^{-3}$  per base pair per division (Berek and Milstein, 1987). Variations in the time at which SHM begins do not affect our results as we start timing AM processes after this point. Each B cell in the dark zone of the GC (the region where this replication/mutation occurs) replicates twice, and we assume that mutation occurs uniformly, neglecting preferential replacements or hypermutation hotspots (Wagner et al., 1995). We assume that the activated B cells are from an appropriate germline (Kepler et al., 2014). The probability of a functionally silent mutation (no change in affinity) is  $p_S = 0.5$ , the probability that a mutation

is lethal (e.g., non-folding) is  $p_L = 0.3$ , and the rest are affinity-affecting mutations (probability  $p_A = 0.2$ ) (Shlomchik et al., 1998). For affinity-affecting mutations, the affinity of a BCR ( $i$ ) with a particular type of antigen ( $j$ ),  $E_{ij}$ , is changed. The extent of the change is chosen from a probability distribution characterized by a long tail of deleterious mutations that reduce affinity, as experimental data suggest that favorable mutations are less likely than deleterious ones (Figure S1A). Our BCR-Ag affinity model is defined in a later sub-section.

### Selection in the Light Zone of the GC

B cells whose re-expressed surface Ig genes do not carry lethal mutations then go through affinity-dependent selection. Two survival signals are required for a B cell to be positively selected. First, the BCR on a B cell must bind to the Ag (immunogen) displayed on the FDC sufficiently strongly to enable internalization. Since processes occur stochastically, in our simulations, B cells internalize Ag with a probability related to the binding affinity of its BCR for the Ag. The greater the difference between the affinity ( $E_{ij}$ ) of a B cell ( $i$ ) for its epitope on Ag  $j$  and a threshold ( $E_a$ ) required for Ag internalization, the greater the probability of Ag internalization. This probability is described mathematically in analogy with a Langmuir isotherm (Equation 1 in Experimental Procedures).

B cells that internalize Ag display antigenic peptides (p) bound to major histocompatibility complex (MHC) molecules on their surface. T cell receptors expressed on T helper cells can bind to these pMHCs to deliver a key survival signal. B cells compete with each other for the limited availability of T cell help (Victoria et al., 2010). T helper cells are specific for peptides derived from the Ag, but they are responsive to diverse pMHCs, not just those derived from the epitope targeted by the BCR. Thus, we assume that the probability of a B cell receiving T cell help is dictated by the amount of internalized Ag, regardless of the identity of its targeted epitope. A B cell that successfully internalizes Ag receives T cell help with a probability that depends upon its probability of internalizing Ag relative to the average probability of internalizing Ag of all the other B cells present in the GC (Equation 2 in Experimental Procedures). A B cell with a high affinity for its target epitope competes better for T cell help.

To study AM against Ag variants that may be present simultaneously, we have to confront an issue for which no experimental information exists: when a GC B cell encounters a FDC, with how many types of Ag variants can it simultaneously interact? We consider two scenarios: (1) each GC B cell interacts with one type of Ag variant during an encounter, or (2) each GC B cell interacts simultaneously with all Ag variants held on the FDCs. We find that which of these scenarios describes the heterogeneity of Ag display on FDCs strongly influences the outcome of AM, indicating the importance of experimental interrogation of this issue.

We study two extreme cases of competition for T cell help: (1) GC B cells compete only with contemporary GC B cells ("peers only") and (2) in addition to current GC B cells, all antibodies generated in previous rounds of mutation/selection also participate in the competition ("Ab feedback"). Reality should be bounded by these scenarios.

### Recycling, Differentiation, and Termination of the GCR

From seminal studies (Oprea and Perelson, 1997), 90% of the selected cells are recirculated to the dark zone. The rest

differentiate into equal numbers of Ab-secreting plasma cells and memory cells that can re-expand upon future activation. Ab feedback is provided by objects in our agent-based model with affinities representing these past GC emigrants. As AM proceeds, at first, the number of B cells in the GC decreases. In some cases, all B cells apoptose and the GC extinguishes. In others, upon reaching a population bottleneck, favorable clones emerge, and the number of B cells rises (Zhang and Shakhnovich, 2010). We terminate the GCR once the GC population recovers to the initial size (i.e., ~1,500 cells), or when an assumed maximum duration of 120 days (or 240 GCR cycles) is reached, whichever comes first. The first condition reflects the fact that an abundance of GC B cells will internalize all the Ag on FDCs. The second condition may reflect antigen decay over time which we do not model explicitly.

#### Model for BCR-Ag Affinity

In classic computational studies focused on GCR stimulated by a single model Ag, B cells were binned into different affinity classes for the Ag (e.g., Kepler and Perelson, 1993). The correlated affinities for different Ag variants require a more detailed description. Other studies have considered mathematical models (such as NK-models) for BCR properties and their Ag affinity that present rugged landscapes for affinity evolution (Deem and Lee, 2003). To unambiguously define affinity to different Ag variants and to consider complex immunogens that reflect features on viral spikes, we developed a coarse-grained model with “residue-level” resolution for key BCR-Ag interactions (Figure 1C). Our model was inspired by the CD4bs on the trimeric HIV-1 viral spike, which is targeted by many monoclonal bnAbs for HIV-1 (Burton et al., 2012). But, it applies to other epitopes that contain highly conserved residues and can easily be modified to consider other pathogens. We ignore distracting epitopes that do not contain conserved residues because they are less likely to elicit bnAbs.

BCRs can potentially make contacts with three types of residues on the viral spike: (1) highly conserved residues of the epitope (18 such residues in model); (2) residues representing motifs, such as glycan attachment sites, which, when occupied, can be associated with sterically blocking access to the conserved residues (6 such residues in model); and (3) variable residues, which upon mutation (includes insertions/deletions), can further mask the conserved residues of the epitope. We include 22 such residues from the variable loops of Env that mimic the most mutable sites in the Seaman neutralization test panel (Seaman et al., 2010; Walker et al., 2011) sequences and publicly available Env sequences.

Each epitope residue is described in a coarse grained manner, such that it is either the wild-type (WT) amino acid or a mutant. The strength of interaction of a residue ( $k$ ) on BCR ( $i$ ) with a residue on the epitope is denoted by  $h_k^i$ . The binding affinity,  $E_{ij}$ , between a B cell clone  $\vec{h}^i$  and a viral strain  $\vec{s}^j$  is modeled as

$$E_{ij}(\vec{h}^i, \vec{s}^j) = \sum_{k=1}^M h_k^i s_k^j + \sum_{k=M+1}^N h_k^i. \quad (1)$$

The first  $M$  interactions (i.e.,  $k \leq M$ ) are with variable contact residues on the viral spike, where  $s_k$  can be either 1 (WT) or -1 (mutated). The other ( $N-M$ ) sites (i.e.,  $k>M$ ) are conserved residues on the epitope with  $s_k = 1$ . The interaction strength  $h_k$  is

drawn from a continuous and uniform distribution within a bounded range (details in “Simulation Methods” in [Extended Experimental Procedures](#)).

Equation 1 is, however, just a starting point for the affinity between a particular BCR and Ag, as BCRs do not interact with peptide chains in a linear fashion, and interactions between residues are not independent, as Equation 1 may imply. Epitope-paratope interactions are distinctly 3-dimensional, and structural aspects of CD4bs bnAbs also point to the importance of how interactions with some residues on the viral spike might influence interactions with other epitope residues (Zhou et al., 2010). For instance, most CD4bs bnAbs avoid contact with almost the entire V1/V2 loop except for a few conserved residues near the stem, and the very potent VRC01 Ab avoids the V5 loop. This suggests that avoiding interactions with some residues can allow better access and stronger interactions with the conserved residues of the epitope. Furthermore, affinity enhancement of bnAbs (such as VRC01) is influenced by alteration of non-contact residues (Klein et al., 2013). We account for these effects as described below (mathematical details in “Simulation Methods” in [Extended Experimental Procedures](#)).

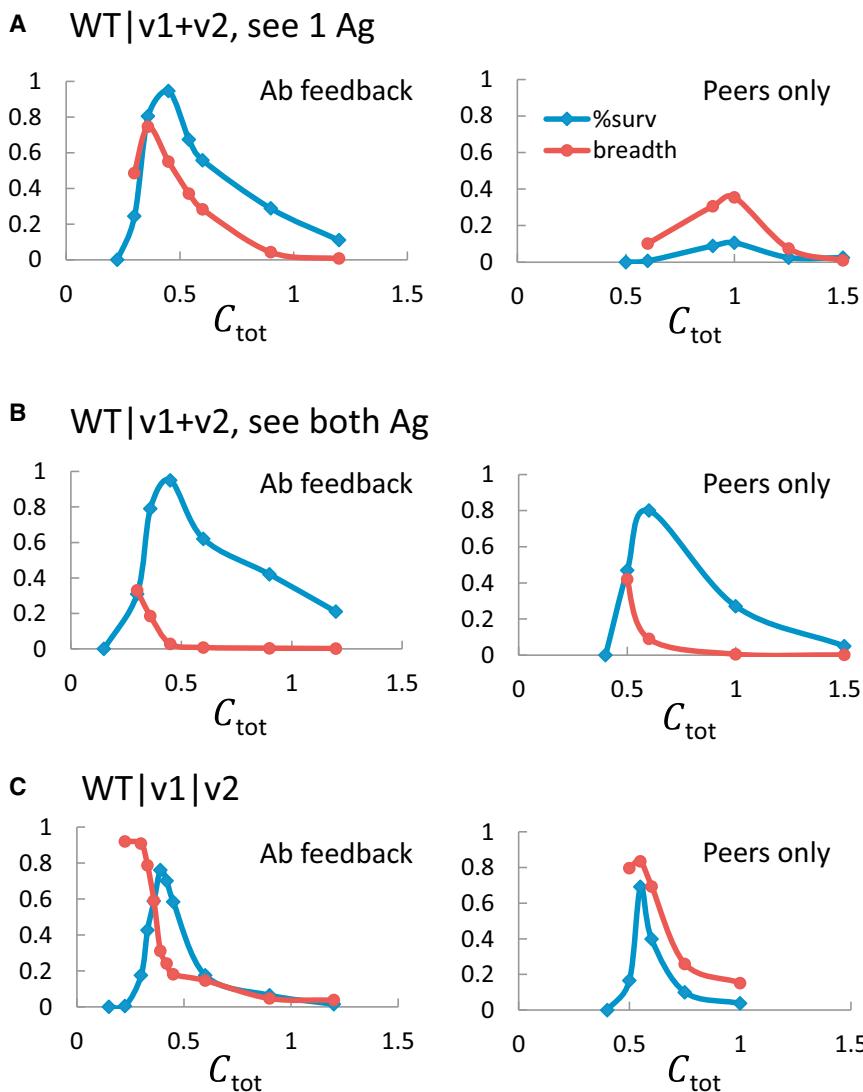
BCR mutations that strengthen (weaken) interactions with the residues that shield the conserved residues could result in a decreased (increased) binding strength (value of  $h$  in Equation 1) for a randomly chosen paratope residue that can potentially interact with a conserved residue on the epitope. This feature reflects the fact that BCRs that decrease contacts with shielding residues are more likely to be able to access and make contacts with the protected conserved residues.

Current knowledge on Env structures (Julien et al., 2013; Lyumkis et al., 2013; Pancera et al., 2014) indicates that V1 is likely to be a dynamic, unfolded, and disordered flexible loop, and mutations in V2, especially insertions/deletions, can hide the conserved residues of a neutralizing epitope. So, paratope alterations that weaken interaction with a mutated variable loop residue result in an increased binding strength (value of  $h$  in Equation 1) for a randomly chosen paratope residue that can potentially interact with a conserved residue on the epitope (Figure 1D), and vice versa.

#### Choice of Immunogens and Immunization Schemes

In silico, we study three Ag variants, the WT Ag (only unmutated residues) and two mutants. Of the 22 most mutable residues in the variable loops, 20 residues are mutated in the majority of the 141 Seaman test panel sequences (Figure S1B). As these highly mutated strains are viable, to maximize the number of non-overlapping mutated residues on the Ag variants, we studied two mutant strains with 11 non-overlapping mutations in the variable sites. We also studied variants with 4 and 8 such mutations. We assume that these Ag variants are not so distal in sequence space that they stimulate completely different B cell lineages.

We investigate three immunization schemes in silico: (1) scheme I (WT+v1+v2): WT Ag and two variants administered as a cocktail. (2) scheme II (WT|v1+v2): Immunization with WT Ag first, followed by administration of the two variants simultaneously. (3) scheme III (WT|v1|v2): Immunization with WT Ag first, followed by sequential administration of the two variants. In our murine experiments we studied schemes I and III.



**Figure 2. Concentration Dependence of GC Survival and Antibody Breadth in “WT First” Schemes**

(A–C) For a relevant range of Ag concentrations ( $C_{\text{tot}}$ ), the fraction of surviving GCs (%surv, blue) and their Ab breadth (red) are shown for (A) scheme II (see 1 Ag), (B) scheme II (see both Ag) and (C) scheme III, with full (left column) or none (right column) Ab feedback. See also Figures S2 and S3.

in 2 weeks, Figure S1D) and the incremental changes in interaction strength (Figure S1E) are consistent with experiment (Berek et al., 1991; Kocks and Rajewsky, 1988; Wedemayer et al., 1997). Higher affinity clones continuously emerge, producing potent Abs (Figure S2 C and F).

#### The Effect of Antigen Concentration

For a single Ag, if Ag dose is too low (Figures S2A–S2B and S2D–S2E, lowest  $C_{\text{wt}}$ ), B cells are unlikely to be selected during GC reactions, the GC collapses, and there is no Ab production. If Ag dose is too high (Figures S2A–S2B and S2D–S2E, highest  $C_{\text{wt}}$ ), selection is easy, there is little competition between B cells, and GCs are rapidly filled with low-affinity clones (Eisen and Siskind, 1964; Goidl et al., 1968). Our agent-based stochastic approach naturally reproduces the inverse correlation between Ag concentration and heterogeneity in Ab affinities (Figures S2G and S2H) observed in early experiments (Eisen and Siskind, 1964).

For each immunization protocol, we simulate many GC reactions. When all steps of immunization and AM are completed, we check whether the affinity of each Ab produced by a GC exceeds a threshold value against each strain in the Seaman panel. Thus, we assess the breadth of cross-reactivity. The breadth of coverage is defined as the fraction of test panel sequences to which an Ab binds with an above-threshold affinity. We collect statistics from many Abs and simulations, and report results as histograms or probabilities of obtaining Abs with a certain breadth. As many GCs are induced in each vaccinated person, these results reflect the probability of obtaining Abs that exhibit particular breadths of coverage in a typical individual.

#### Model Reproduces Known Features of Affinity Maturation with a Single Antigen

Immunization with only the WT Ag leads to affinity enhancement on conserved and variable residues alike (Figure S1E). The number of accumulated mutations (seven affinity-affecting mutations

For the three immunization schemes with antigen variants, Ag concentration is again a very important variable (Figure 2). If the Ag concentration is too low, most GCs collapse as expected. However, there is a curious decline in GC survival and Ab breadth if the Ag concentration is too high for immunization schemes II and III. During AM after immunizing with a high dose of WT Ag, many B cells survive easily and the GC reaction draws to an end quickly. Therefore, these B cells are largely low-affinity clones that have not accumulated mutations that enhance contacts with the conserved residues of the epitope. So, upon administering either a cocktail of the two Ag variants or just one variant, these B cells have a small chance of binding sufficiently strongly with the mutants and surviving, and so GCs collapse despite the abundance of Ag. The few Abs produced do not develop breadth as the large number of mutations required to confer breadth do not evolve (see examples in Figure S3).

Hereon, for each immunization scheme, we show results for the Ag concentration (listed in Table 1) that yields Abs with a

**Table 1. Properties of Antibody Responses Produced by Various Immunization Schemes**

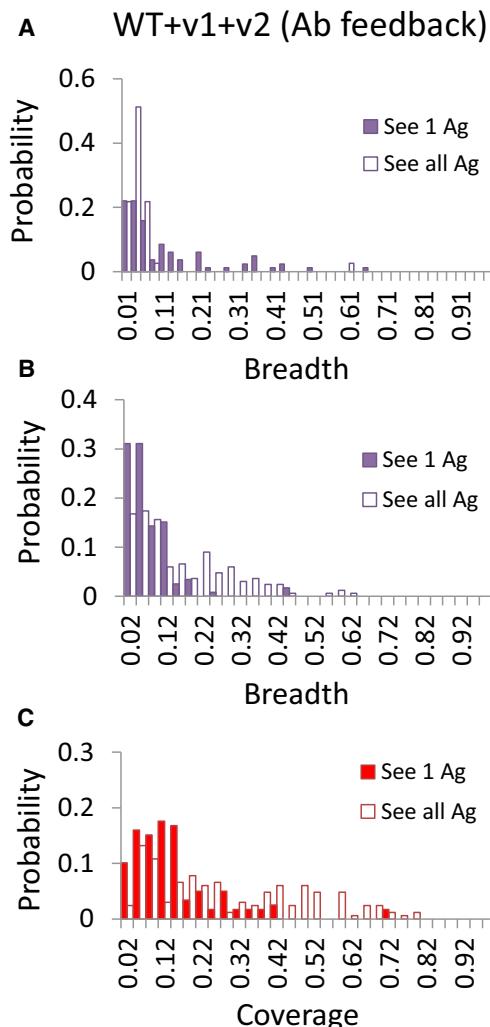
Scheme	%surv	Breadth	$h_c$	$h_v$	$C_{tot}$
<b>WT only (<math>t_{max} = 80</math>)</b>					
Ab feedback	54%	0.6%	0.45	0.39	0.3
peers only	79%	0.7%	0.49	0.43	0.5
<b>(I) WT+v1+v2 (<math>t_{max} = 240</math>)</b>					
See 1 Ag; Ab feedback	1%	11%	0.72	0.29	0.6
peers only	1%	2%	0.52	0.22	1.0
See all Ag; Ab feedback	1%	5%	0.84	0.70	0.3
peers only	3%	3%	0.89	0.83	0.5
<b>(II) WT v1+v2 (<math>t_{max} = 80 + 160 = 240</math>)</b>					
See 1 Ag; Ab feedback	80%	75%	1.21	0.16	0.36
peers only	10%	35%	0.73	0.16	1.0
See both Ag; Ab feedback	31%	33%	1.07	0.37	0.3
peers only	47%	42%	1.22	0.41	0.5
<b>(III) WT v1 v2 (<math>t_{max} = 80 \times 3 = 240</math>)</b>					
Ab feedback	43%	79%	1.38	0.22	0.33
peers only	69%	83%	1.64	0.24	0.55

%surv: percentage of seeded GCs that produce antibodies efficiently; Breadth: defined in text;  $h_c$  ( $h_v$ ): the average strength of interactions with conserved (variable) residues of the epitope—large values of  $h_c$  indicate focusing on the conserved residues of the epitope;  $C_{tot}$ : antigen concentration.

near optimal breadth (peak of red curves in Figure 2). In scheme III, the chosen concentration is slightly higher to ensure high Ab production.

### Immunizing with a Cocktail of WT Ag and Two Variants Fails to Elicit an Effective Response

Figure 3A shows *in silico* results for the distribution of the breadth of coverage (for the Seeman panel) of Abs produced using immunization scheme I (cocktail of WT+v1+v2). Ag concentration is fixed to yield the greatest breadth, but two more variables could be important: during selection in the light zone, does a B cell-FDC encounter involve one or all Ag variants? Is “Ab feedback” important during AM? For this immunization scheme, in most scenarios, just a few percent of GCs survive (Table 1); i.e., Ab titers are predicted to be low. To see why, consider first the situation where B cells interact with only one Ag variant during each encounter with FDCs. B cells that bind moderately strongly to one of the three variants seed the GC, and they do not have any strong interactions with the conserved residues of the epitope yet. Now consider a B cell that is selected by a particular Ag variant in an early round, then mutates, and returns to the light zone to encounter a different variant. Without even moderately strong contact with the conserved residues, the probability that such a B cell is cross-reactive to the new variant is small. Thus, apoptosis is the most likely outcome, leading to collapse of the GC. For the rare GCs that do survive, there is evolution of interactions with the conserved residues and diminution of interactions with variable loops during AM, and this confers some breadth (relative to immunization with WT Ag alone, see Table 1).

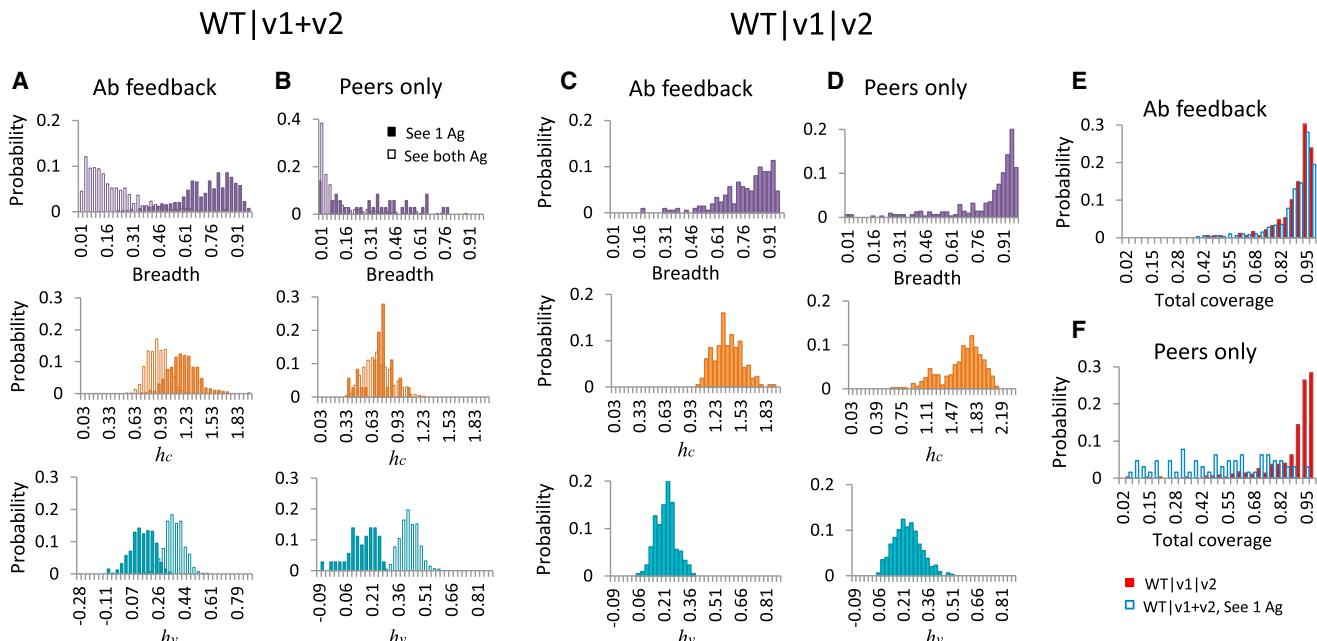
**Figure 3. Distribution of Breadth of Individual Antibodies and Their Total Coverage for Surviving GCs in Scheme I—WT+v1+v2**

(A and B) Histograms of breadth for having 11 (A) or 4 (B) non-overlapping mutations in the two variants. Histograms are obtained as a distribution for all the surviving GCs.

(C) Histograms of coverage for surviving GCs under the same conditions as in (B). The coverage is defined as a sum of non-overlapping specificities. Results are shown for the cases of “see 1 Ag” (filled bars) and “see all Ag” (unfilled bars), with Ab feedback.

See also Figure S4.

If B cells interact with all Ag variants during each encounter with FDCs, and the Ag concentration is chosen to be the one that confers optimal breadth, the percentage of surviving GCs and the distribution of breadth is not very different compared to the case where only one Ag variant is encountered at a time (Figure 3A). However, importantly, the optimal concentration is lower when B cells can interact with all Ag variants. If the Ag concentrations were made the same as when B cells interact with one Ag at a time, then most GCs would survive when all Ag variants are simultaneously encountered (Figure S4B). This is because, in this case, B cells that bind sufficiently strongly to



**Figure 4. Comparison of Statistics for Surviving GCs in Scheme II—WT|v1+v2 and Scheme III—WT|v1|v2**

(A–F) Shown are histograms for breadth, mean interaction strength with conserved residues ( $h_c$ ) and with variable residues ( $h_v$ ) (A–D), and for total coverage (E and F) via a polyclonal response. Results are shown for “Ab feedback” (A, C, and E) and “peers only” (B, D, and F) scenarios. For scheme II (A and B), cases of “see 1 Ag” (filled bars) and “see both Ag” (unfilled bars) are shown. Total coverage (E and F) is presented for scheme III (WT|v1|v2) by red filled bars and for scheme II (WT|v1+v2, see 1 Ag) by blue unfilled bars. See also Figure S5.

any one Ag variant can progress through AM. As a consequence, the generated Abs are likely to be strain specific. Thus, if Ag is displayed homogeneously on FDCs, for immunization scheme I, optimal conditions for inducing Abs efficiently at high titers would not induce bnAbs. The lower Ag concentration required for optimal breadth results in the survival of only a relatively low fraction of GCs (Figure 3A).

We also carried out calculations where the Ag variants differ from each other by only four non-overlapping mutations. When all variants are simultaneously encountered in each interaction with FDCs and Ab feedback is turned on, comparing results for this case (Figure 3B) with that when the variants differ by 11 mutations (Figure 3A), we find that a much higher fraction (10% versus 1%) of GCs succeed in producing Abs (Figure S4D). This is because the Ags vary less from each other, and so the chance of a B cell being cross-reactive to variants in early stages of AM is higher, thus increasing the probability of it being positively selected. The breadth of the Abs produced is not large because many variable residues that are unmutated in the Ag variants are mutated in the test panel sequences, and strain specific Abs are likely to be produced as all Ag variants on FDCs are simultaneously encountered. This is why the total coverage (Figure 3C), defined as the breadth of coverage offered by the polyclonal response from all the Abs, is higher than the average breadth of coverage of individual Abs (Figure 3B). Results for the case where only one Ag variant is encountered at a time on FDCs are shown in Figures 3B and 3C (filled bars), and for no Ab feedback see Table 1.

In our in silico studies of this immunization scheme, we did not find any conditions that result in bnAbs with high probability (Figure S4). This is because B cell evolution during AM is made difficult by the potentially conflicting selective pressures imposed by Ag variants simultaneously displayed on FDCs. We term this phenomenon, *frustrated AM*. The degree to which AM is frustrated is controlled by a complex interplay between Ag concentration, heterogeneity of Ag display on the FDCs, the number of mutations that separate the Ag variants, and the extent to which Ab feedback is important.

#### Immunizing with WT Ag First and then a Cocktail of the Two Variants Yields Abs with Significant Breadth in Special Circumstances

Consider first the situation in which there is no Ab feedback and only one Ag variant is encountered at a time on FDCs (Figure 4B filled bars). In the first period of AM induced by the WT Ag, moderately strong contacts with the conserved residues of the epitope can evolve. During AM following immunization with the cocktail, individual B cells have equal chance of encountering either Ag variant during selection. Since the two variants have non-overlapping mutated residues, beneficial BCR mutations for one strain are deleterious for the other. Maturing B cell lineages are frustrated in satisfying these conflicting requirements for selection in successive rounds, but less so than in scheme I because moderately strong interactions with the conserved residues have evolved during AM driven by the WT Ag prime, resulting in some cross-reactivity. Therefore, a greater fraction (~10%) of GCs survive compared to scheme I. Since a relatively high Ag

concentration is required for GC survival in this case (Figure 2A right panel blue curve), B cells can expand readily and GCRs terminate quickly. Short maturation times limit the average breadth (~35%, Figure 2A right, red curve), but Abs with large breadth do evolve with low probability.

If there is Ab feedback, lower Ag concentrations still allow GCs to mature successfully. This is because Abs produced in previous rounds of AM often have lower affinities for the encountered Ag than the best clones produced during the current round of mutation and selection. This confers a competitive advantage to these best B cell clones because they are not just competing with each other, but also with the weaker affinity Abs ("Simulation Methods" in [Extended Experimental Procedures](#)). Also, the Abs compete effectively with unfavorable mutant cells that may emerge and be stochastically selected only to be extinguished in a future round of mutation-selection. Lowering Ag concentration diminishes the probability of selection in each round, thus allowing AM to proceed for a longer time before all the Ag on FDCs is internalized. This allows mutations that confer breadth to accumulate. This balance of lowering frustration through Ab feedback and increasing frustration by reducing Ag concentration results in a very high survival rate for GCs and considerable breadth in the Ab response (Figure 2A left, Figure 4A filled bars).

If multiple Ags are encountered during each interaction with FDCs, the outcome is different (Figures 4A and 4B, unfilled bars). Since a choice of being selected by the same variant (rather than a randomly chosen one of the two) is always available in successive rounds of AM, the frustration is low. In most cases, we find that B cells are selected by one of the variants repeatedly, and so strain-specific Abs are likely to evolve (Table 1). Furthermore, access to both Ag variants allows B cells to multiply successfully in the GC and quickly consume all the Ag on FDCs. The duration of AM is short, so mutations that confer breadth evolve rarely, and the average breadth is no more than 40%, regardless of the extent of Ab feedback (Table 1). For this optimal breadth to develop, the antigen concentrations must be much lower than when only one Ag variant is seen during each B cell-FDC encounter (Figure 2B versus 2A).

Our results suggest that this immunization scheme has the potential to produce bnAbs (~75% breadth) with high probability if B cells encounter one Ag variant at a time on FDCs and all Abs produced in previous rounds migrate through ongoing GCs. Rather than the extreme situations of full Ab feedback or none at all, there is some middle ground, and also most likely, B cells encounter one FDC Ag variant in some rounds of selection and all variants during others. Thus, our results suggest that, depending upon circumstances, this immunization scheme produces Abs efficiently with a probability between 10%–80%, and with average breadths ranging from 30% to 80% (Table 1). This scheme may be analogous to what ensues in patients who develop bnAbs upon natural infection as they are first exposed to the infecting strain, which then diversifies. However, the sensitivity of our results to varying conditions suggest that it may be quite difficult to induce bnAbs consistently in diverse patients using this immunization scheme. This is because the conditions noted above are likely to vary between and within individuals.

### Sequential Immunization with Antigen Variants Leads to Efficient Induction of an Antibody Response with Broad Specificity

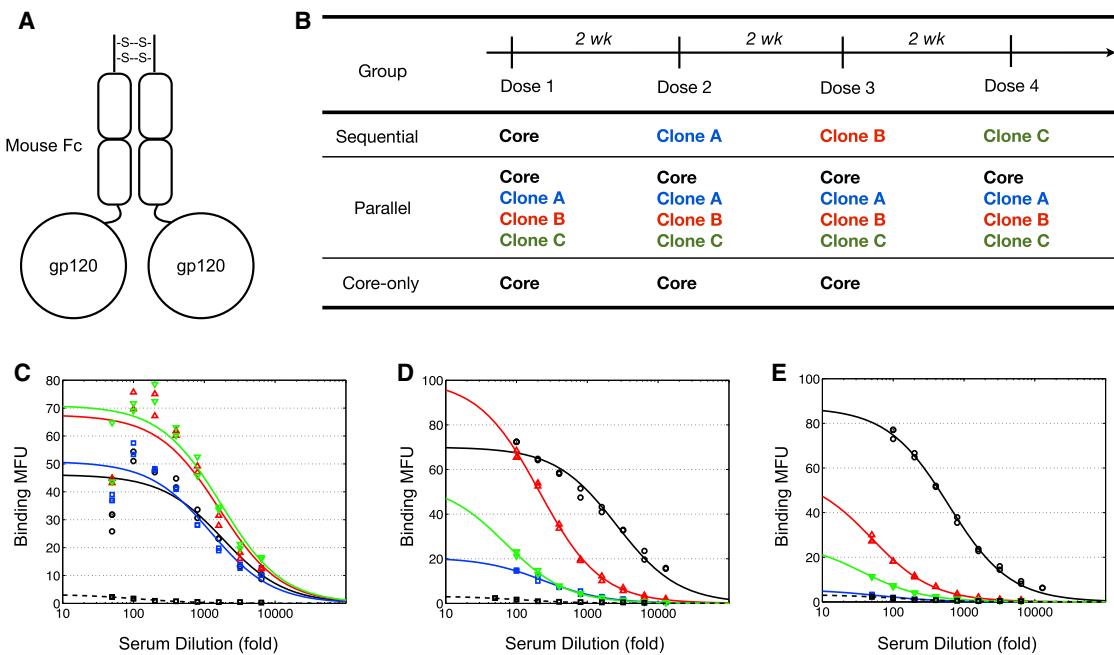
Sequential immunization with the three Ag variants temporally separates the mutually conflicting selective driving forces imposed by multiple variants. Confounding effects associated with whether one or multiple types of Ags participate in each B cell-FDC encounter in the GC are also obviated.

Consider first the situation where there is no Ab feedback (Figure 4D). Maturation against the WT Ag results in some strong interactions with the conserved residues. When AM ensues with the first variant, two main types of lineages evolve in the simulations (Figure S5). Both types of lineages tend to reduce interactions with the residues that are mutated in the first variant Ag. But, in one type of lineage, interactions with the unmutated variable residues in the first variant (which are mutated in the second variant) are enhanced; in the other type of lineage, these interactions do not change much. When the second variant with non-overlapping mutations is introduced, the latter lineages have a good chance of outcompeting the former ones during AM. This is because clones that did not increase their footprint on the variable residues that are now mutated in the second variant can simply reduce interactions with these residues, and focus strongly only on the conserved residues of the epitope. This is why Abs produced in scheme III exhibit a narrow distribution of very large breadth (Figure 4D top panel) rather than a broad distribution of Abs with moderate breadth seen in the analogous situation using scheme II (Figure 4B top panel, filled bars). In scheme II, lineages specific for one or the other variant strain have a higher chance of survival. In scheme III, Abs that focus contacts with the conserved residues of the epitope and minimize interactions with all the variable residues are more likely to emerge (Figures 4B and 4D, middle panels).

Comparing "Ab feedback" (Figure 4C) and "peers only" scenarios (Figure 4D) in scheme III, we see that Ab feedback leads to narrower distributions of Abs with large breadth, due to the same effects of Ab feedback noted for scheme II. As shown in Figures 4E and 4F, the distributions of total coverage for Abs produced by schemes II and III are very close when there is Ab feedback (Figure 4E), and the average values are both as high as 90%. If there is no Ab feedback (Figure 4F), however, scheme III gives a similar broad coverage (89%), whereas scheme II yields a moderate coverage (51%) even with polyclonal responses. So, scheme III is predicted to be more robust to many varying conditions.

### Experiments in Mice with a Model System Show that Antibodies that Focus on the Conserved Residues Emerge upon Sequential Immunization, but Not upon Immunization with a Cocktail

Our calculations predict that, compared to the other immunization schemes we studied, truly cross-reactive Abs that focus contacts on the conserved residues of the epitope with a small footprint outside are most likely to emerge robustly upon sequential immunization with Ag variants. Immunization with a cocktail of Ags is predicted to result in a significantly worse outcome compared to sequential immunization. We



**Figure 5. Immunization of Mice with Variant Immunogens**

(A) Schematic of an Fc-gp120 immunogen.

(B) Immunization groups and dosing schedule for mouse experiment.

(C–E) Serum titrations (day 48) of a representative mouse from each group—(C) sequential, (D) parallel, (E) core-only—on yeast displaying stripped core (black), clone A (blue), clone B (red), or clone C (green). Each plot also includes serum from an unimmunized mouse binding to yeast displaying clone C (black dashed). Binding data are fit to a monovalent binding isotherm of the form  $y = y_{max} \text{dil}^{-1} / (\text{dil}^{-1} + K_A^{-1})$  where “dil” is the serum dilution and  $y$  is the binding signal in MFU. Discussion of the fitted parameters can be found in the supplemental material.

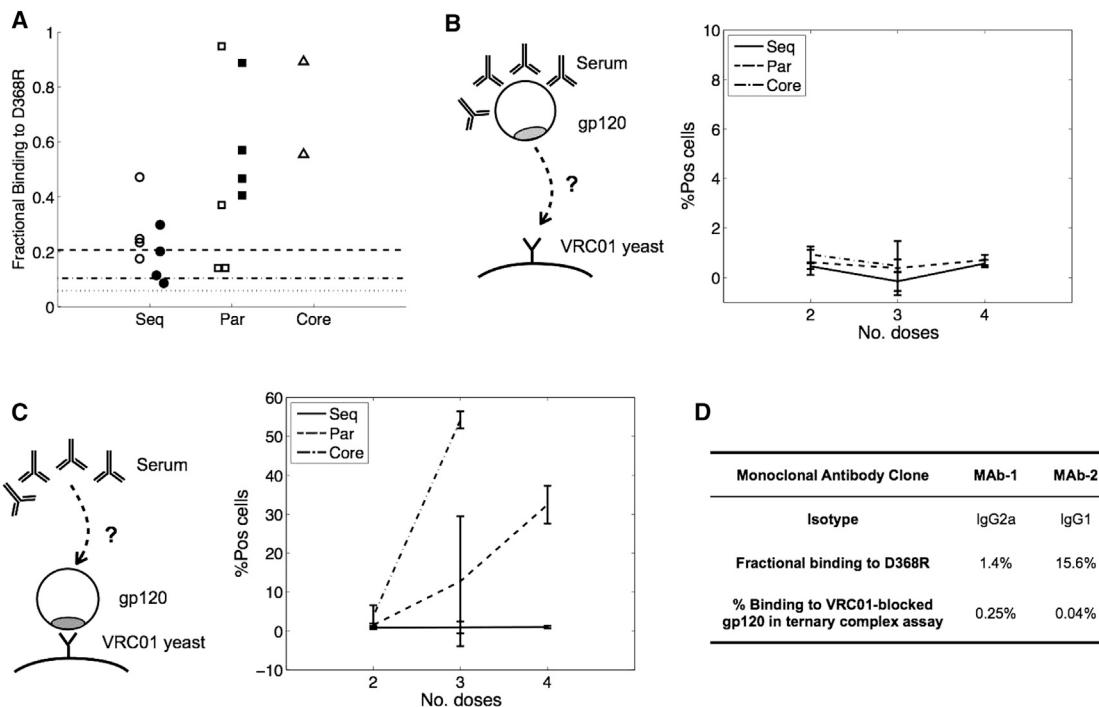
See also Figure S6 and Table S1.

note previous immunization studies that have suggested the advantages of sequential immunization of HIV Env in generating cross-reactive Abs, in further agreement with our computational results (Malherbe et al., 2011; Pissani et al., 2012). Here, we tested our predictions comparing sequential immunization and cocktails in mouse studies using precisely engineered Ag variants. It is important to note that the variant Ags we have engineered are gp120 monomers of Env, not trimers. Thus, they cannot be expected to produce bnAbs that neutralize virus particles with intact trimeric spikes. The goal of our experiments was to precisely test whether cross-reactive Abs that focus on the conserved residues of an epitope (CD4bs) are more likely to develop upon sequential immunization, compared to administration of a cocktail of the same immunogens, as per our *in silico* predictions.

As described in [Extended Experimental Procedures](#), we engineered variants of a minimized gp120 core protein immunogen (the “stripped core”) that incorporate diversification at 43 surface residues outside the CD4bs, while retaining binding affinity to the VRC01 bnAb ([Table S1](#), [Figure S6](#)). These immunogens are designed to present biochemically novel residues outside the CD4bs such that only the desired CD4bs epitope is conserved across all the immunogens. We used four engineered variants (including stripped core) that fit these criteria. Our *in silico* results should hold irrespective of whether three or four variants are used.

BALB/c mice were immunized intranasally with HEK293-produced immunogens fused to mouse Fc ([Figure 5A](#)) with a CpG oligonucleotide adjuvant, following protocols described elsewhere ([Ye et al., 2011](#)). Immunogens were administered every 2 weeks at a total protein dose of 50 pmol. Animals were divided into three groups ([Figure 5B](#)). The “Sequential” group of four mice was immunized sequentially with each of the four Ag variants. The “Parallel” group of four mice received a cocktail of all four variants at each administration. The “core-only” group of two mice was given three doses of the same stripped core immunogen. Serum was collected every week.

Serum binding to the Ag variants was assayed by flow cytometry of yeast displaying each of the four variants. Representative curves for each of the three immunization groups are shown in [Figures 5C–5E](#). Sequentially immunized mice exhibit similar binding to each of the four Ag variants, consistent with serum that recognizes and focuses contacts with a shared epitope presented on all variants ([Figure 5C](#)). Mice immunized with a cocktail of four variants show a broader spread of binding affinities, consistent with dominant serum specificities for some variants but not others ([Figure 5D](#)). In the mouse represented, for example, it appears that the serum recognizes the epitope on stripped core that is not present on clones A or C. Serum from mice immunized only with stripped core bind strongly to stripped core but not to the other immunogens, consistent with our immunogen design objective that the four variants not share common



**Figure 6. Determination of Serum Specificity for the CD4 Binding Site**

(A) Fractional binding to D368R mutant versus stripped core of serum from each group after three (open markers) and four (filled markers) immunizations. The fractional binding to D368R for antibody VRC01 at various concentrations is shown as lines: 0.21 at 93 nM (dashed line), 0.10 at 75 nM (dash-dotted line), and 0.06 at 8.2 nM (dotted line).

(B) Schematic and data from ternary complex assay in which gp120 (CD4bs shown in gray) is pre-incubated with serum then added to VRC01 yeast. The formation of a ternary complex is only possible if VRC01 is able to bind to its epitope on gp120 after serum has bound. The figure shows the percentage of yeast cells positive for mouse serum from each immunization group—sequential (solid line), parallel (dashed line), core-only (dash-dotted line)—for various time points (mean  $\pm$  SD).

(C) Schematic and data from ternary complex assay in which gp120 is pre-loaded onto VRC01 yeast then incubated with serum. The formation of a ternary complex is only possible if serum antibodies are able to bind to their epitope(s) after VRC01 has bound. The figure shows the percentage of yeast cells positive for mouse serum from each immunization group as above.

(D) Analysis of two monoclonal antibodies from a sequentially immunized mouse: fractional binding to D368R versus stripped core gp120 was assayed with 10 nM monoclonal antibody (analogous to 6A); percentage binding in the ternary complex assay (analogous to 6C), in which gp120 is pre-bound to VRC01 yeast.

surface residues outside the CD4bs (Figure 5E), like the non-overlapping mutations in Ag variants in our calculations.

Serum specificity for the CD4bs was determined in two ways. First, we measured the binding of serum to the stripped core mutant D368R, which is known to disrupt the binding of several CD4bs-directed Abs (Thali et al., 1992). Serum samples collected after the third and fourth immunizations were incubated at a 1:100 dilution with yeast displaying either stripped core or the D368R mutant (Figure 6A). For mice from the sequential immunization group, the fractional binding of serum to D368R was similar to that of VRC01 at saturating concentrations. Of the mice that were immunized with a cocktail of all four immunogens, three are very sensitive to D368R after three immunizations but appear to lose this specificity by the fourth administration.

A second assay used to determine specificity was to evaluate the simultaneous binding of VRC01 and serum antibodies to gp120 (Experimental Procedures). The assay was performed first by pre-incubating the gp120 with mouse serum before introducing VRC01-displaying yeast (Figure 6B). If serum antibodies compete with VRC01 for its binding site, the gp120-serum com-

plex will not bind to the yeast. No binding signal was observed for any of the serum samples, suggesting that all contained some fraction of VRC01-competitive antibodies. This result is consistent with the D368R gp120 binding results discussed above.

The same assay was then performed by pre-incubating gp120 with VRC01 yeast before introducing the serum (Figure 6C) to determine whether the serum contains any specificity other than to the CD4bs. Note that due to plasmid loss in yeast during culture, ~60% positive cells is the maximum observable signal in the assay. Mice immunized with just stripped core (53%, 56%) or with a cocktail of four Ags (mean = 32%  $\pm$  5%) developed antibodies that bind to other epitopes on gp120. However, mice immunized sequentially did not (mean = 1.0%  $\pm$  0.3%). These experiments suggest that sequential immunization with Ag variants elicited a serum response that was VRC01-competitive and entirely focused on the conserved residues targeted by the VRC01 Ab. Mice that were exposed to a cocktail of the variants generated some VRC01-competitive antibodies, but also Abs that targeted non-conserved residues on a single immunogen, i.e., were not cross-reactive. In other words, only sequential

immunization robustly elicited a response that is focused on the conserved residues of the desired epitope. These findings are in harmony with our *in silico* results.

To establish that individual CD4bs-competitive antibodies were generated by sequential immunization, hybridomas were generated by splenotypic fusion from a mouse immunized using this protocol. Two distinct gp120-binding monoclonal antibodies were isolated (for sequences, see [Extended Experimental Procedures](#)). The antibodies both targeted the CD4bs epitope, as both were disrupted by the D368R mutation (at 10 nM antibody, fractional binding to D368R versus stripped core gp120 of 1.4% and 15.6%) and neither was able to bind to stripped core that had been pre-blocked with VRC01 ([Figure 6D](#)). These data make clear that the monoclonal antibodies are indeed focused on the CD4bs upon sequential immunization with the variant antigens.

## DISCUSSION

The induction of bnAbs against highly mutable pathogens, such as HIV and influenza, will require immunization with Ag variants. Rational design of immunogens and efficient immunization protocols that elicit bnAbs requires mechanistic understanding of a basic problem in immunobiology: viz., how AM occurs in the presence of variant Ags. Our studies of AM in this setting have revealed new concepts, suggested new avenues for experimental and theoretical research, and have practical value. An important concept that emerges is that, when multiple Ag variants are used as immunogens, they can present conflicting selection forces during AM which frustrates the evolution of Abs. This phenomenon is most acutely manifested upon immunization with a cocktail of Ag variants. In this case, a B cell of a particular lineage that is selected in one round by a particular Ag variant is likely, after the next round of mutation, to encounter a different Ag variant on FDCs. This B cell is unlikely to be positively selected in this encounter, thus ending the evolution of a potentially favorable B cell lineage. The degree to which AM is frustrated is determined by a complex interplay of effects dependent on Ag concentration, how heterogeneously Ag variants are displayed on FDCs, the number of mutations that separate the Ag variants, and the extent to which previously generated Abs migrate through ongoing GCs. Our results provide empirical evidence that these parameters can be tuned to achieve an optimal level of frustration. Too much frustration prevents GC reactions from evolving favorable mutations that confer breadth. Too low a level of frustration (the extreme case being immunization with one Ag) results in the development of strain-specific antibodies. An avenue for future theoretical research is to define a quantitative metric of frustration that describes precisely how the degree of frustration depends on the pertinent variables. The metric of frustration could then be optimized to design effective immunization protocols.

Our *in silico* results suggest that if one primes with the WT Ag, and then boosts with a cocktail of variants, bnAbs can emerge but only if some special conditions hold. Such a scenario may have been the driver of bnAb production in naturally infected patients because this immunization strategy mimics

natural infection with one strain and subsequent viral diversification. However, our results show that deviations from these special conditions make the evolution of bnAbs unlikely in this setting ([Table 1](#)). This suggests that potent bnAbs evolve during natural infection only when special conditions are met, which happens rarely, and thus take a long time to emerge by chance. Controlling immunization protocols precisely to satisfy special conditions in every vaccinated individual seems difficult. Furthermore, even if the conditions were set to perfectly mimic a situation that resulted in the evolution of bnAbs upon natural infection, the outcome may not be favorable because Darwinian evolution is inherently stochastic and cannot be replicated exactly. So, although it is important to understand how bnAbs are produced during natural infection, mimicking these conditions may not be an efficient way to induce bnAbs by vaccination.

We find that when Ag variants are administered sequentially, bnAbs can develop with relatively high probability over a wider range of conditions than if other strategies are deployed ([Table 1](#), and [Figures 2, 3, and 4](#)). In this case, the conflicting selection forces due to the variant Ags are temporally separated. This can lead to successful evolution of multiple B cell lineages that first develop moderate interactions with the conserved residues of the epitope and then progressively acquire mutations that evade contacts with one set of variable residues and then another. Evolution of a set of possible clones which have a chance at maturing into bnAbs after AM against the first variant makes success more likely as all of them are unlikely to go extinct during stochastic mutation and selection upon immunizing with the second variant. Our *in silico* results predict that this immunization strategy is superior to immunization with a cocktail of the same set of variants, and our model experiments in mice support this prediction.

Sequential immunization with Ag variants that share a single epitope to induce B cells specific for the common epitope has been tried with linear peptide epitopes derived from HIV gp41, though the resulting Abs were non-neutralizing ([Guenaga et al., 2011; Correia et al., 2010](#)). Past work also emphasizes the importance of masking irrelevant epitopes, but by itself, this strategy is insufficient for successful vaccination ([Pantophlet et al., 2003; Selvarajah et al., 2005](#)). Our *in silico* and *in vivo* studies suggest that precisely engineered variant intact virus spike immunogens administered in a sequential fashion under appropriate conditions may offer the best opportunities for induction of bnAbs against HIV-1 and other highly mutable pathogens.

## EXPERIMENTAL PROCEDURES

### In Silico Model

The probability of a B cell internalizing antigen is given by:

$$P_a^i = \frac{\sum_{j=1}^{n_A} C_j e^{(E_j - E_a)/k_B T}}{1 + \sum_{j=1}^{n_A} C_j e^{(E_j - E_a)/k_B T}}. \quad (1)$$

Here  $C_j$  is the concentration of Ag of type  $j$  presented on the FDCs. The sum over the Ag index  $j$  runs through  $n_A$  distinct types of Ags that B cell  $i$  could potentially interact with simultaneously during an encounter with a FDC bearing multiple Ag variants.

The probability of a B cell to succeed in receiving T cell help is given by:

$$P_{Th}^i = \frac{\sum_{j=1}^{n_A} e^{E_{ij}/k_B T}}{\sum_{j=1}^{n_A} e^{E_{ij}/k_B T} + C_{tot}^{-1} \left\langle \sum_{j'=1}^{n_A} e^{E_{ij'}/k_B T} \right\rangle_{i' \neq i}}, \quad (2)$$

where  $\left\langle \sum_{j'=1}^{n_A} e^{E_{ij'}/k_B T} \right\rangle_{i' \neq i}$  is the average probability of internalizing Ag of all the other B cells present in the GC.  $C_{tot} = \sum_{j=1}^{n_A} C_j$  is the total Ag concentration, and the dependence on this variable accounts for the observation that the number of activated T helper cells increases with Ag dose, thus making it more likely that B cells receive T cell help.

#### Binding Assay to Determine Competitiveness with VRC01

Yeast displaying the scFv of VRC01 was incubated with soluble stripped core gp120 and an excess of mouse antiserum. The formation of a ternary complex—a sandwich of VRC01, gp120, and mouse serum—was detected as the presence of mouse antibodies on yeast by flow cytometry (% positive cells). The experiment was done in two ways as described in text.

See [Extended Experimental Procedures](#) for additional information.

#### SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, six figures, and one table and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2015.01.027>.

#### AUTHOR CONTRIBUTIONS

S.W., A.K.C., M.K., and D.R.B. conceived and designed in silico studies; S.W., A.K.C., M.K. analyzed in silico data; S.W. performed in silico studies; J.M.F., D.J.I., M.H., K.D.W. designed experiments and J.M.F., K.D.W. analyzed data; J.M.F. performed experiments; B.K. helped engineer Ag variants; A.K.C., S.W., M.K., J.M.F., K.D.W., D.R.B., H.N.E. wrote paper.

#### ACKNOWLEDGMENTS

Financial support provided by the Ragon Institute of MGH, MIT, and Harvard (AKC, SW, MK, KDW, JMF) and the International AIDS Vaccine Initiative (IAVI) through the Neutralizing Antibody Consortium SFP1849 (DRB); NIH grants R01 AI033292 (DRB), and Center for HIV/AIDS Vaccine Immunology and Immunogen Discovery grant UM1AI100663 (DRB, AKC). We are grateful to Kevin Kaczorowski and Dariusz Murakowski for helpful discussions and comments.

Received: July 2, 2014

Revised: October 3, 2014

Accepted: December 19, 2014

Published: February 5, 2015

#### REFERENCES

Allen, C.D.C., Okada, T., Tang, H.L., and Cyster, J.G. (2007). Imaging of germinal center selection events during affinity maturation. *Science* 315, 528–531.

Batista, F.D., and Neuberger, M.S. (1998). Affinity dependence of the B cell response to antigen: a threshold, a ceiling, and the importance of off-rate. *Immunity* 8, 751–759.

Baumjohann, D., Preite, S., Reboldi, A., Ronchi, F., Ansel, K.M., Lanzavecchia, A., and Sallusto, F. (2013). Persistent antigen and germinal center B cells sustain T follicular helper cell responses and phenotype. *Immunity* 38, 596–605.

Berek, C., and Milstein, C. (1987). Mutation drift and repertoire shift in the maturation of the immune response. *Immunol. Rev.* 96, 23–41.

Berek, C., Berger, A., and Apel, M. (1991). Maturation of the immune response in germinal centers. *Cell* 67, 1121–1129.

Burton, D.R., Ahmed, R., Barouch, D.H., Butera, S.T., Crotty, S., Godzik, A., Kaufmann, D.E., McElrath, M.J., Nussenzweig, M.C., Pulendran, B., et al. (2012). A blueprint for HIV vaccine discovery. *Cell Host Microbe* 12, 396–407.

Correia, B.E., Ban, Y.E., Holmes, M.A., Xu, H., Ellingson, K., Kraft, Z., Carrico, C., Boni, E., Sather, D.N., Zenobia, C., et al. (2010). Computational design of epitope-scaffolds allows induction of antibodies specific for a poorly immunogenic HIV vaccine epitope. *Structure* 18, 1116–1126.

Deem, M.W., and Lee, H.Y. (2003). Sequence space localization in the immune system response to vaccination and disease. *Phys. Rev. Lett.* 91, 068101.

Eisen, H.N., and Siskind, G.W. (1964). Variations in affinities of antibodies during the immune response. *Biochemistry* 3, 996–1008.

Gao, F., Bonsignori, M., Liao, H.X., Kumar, A., Xia, S.M., Lu, X., Cai, F., Hwang, K.K., Song, H., Zhou, T., et al. (2014). Cooperation of B cell lineages in induction of HIV-1-broadly neutralizing antibodies. *Cell* 158, 481–491.

Goidl, E.A., Paul, W.E., Siskind, G.W., and Benacerraf, B. (1968). The effect of antigen dose and time after immunization on the amount and affinity of anti-hapten antibody. *J. Immunol.* 100, 371–375.

Guenaga, J., Dosenovic, P., Ofek, G., Baker, D., Schief, W.R., Kwong, P.D., Karlsson Hedestam, G.B., and Wyatt, R.T. (2011). Heterologous epitope-scaffold prime: boosting immuno-focuses B cell responses to the HIV-1 gp41 2F5 neutralization determinant. *PLoS ONE* 6, e16074.

Jacob, J., Kassir, R., and Kelsoe, G. (1991). In situ studies of the primary immune response to (4-hydroxy-3-nitrophenyl)acetyl. I. The architecture and dynamics of responding cell populations. *J. Exp. Med.* 173, 1165–1175.

Julien, J.-P., Cupo, A., Sok, D., Stanfield, R.L., Lyumkis, D., Deller, M.C., Klasse, P.J., Burton, D.R., Sanders, R.W., Moore, J.P., et al. (2013). Crystal structure of a soluble cleaved HIV-1 envelope trimer. *Science* 342, 1477–1483.

Källberg, E., Jainandunsing, S., Gray, D., and Leanderson, T. (1996). Somatic mutation of immunoglobulin V genes in vitro. *Science* 271, 1285–1289.

Kelsoe, G. (1996). The germinal center: a crucible for lymphocyte selection. *Semin. Immunol.* 8, 179–184.

Kepler, T.B., and Perelson, A.S. (1993). Somatic hypermutation in B cells: an optimal control treatment. *J. Theor. Biol.* 164, 37–64.

Kepler, T.B., Munshaw, S., Wiehe, K., Zhang, R., Yu, J.S., Woods, C.W., Denny, T.N., Tomaras, G.D., Alam, S.M., Moody, M.A., et al. (2014). Reconstructing a B-cell clonal lineage. II. mutation, selection, and affinity maturation. *Front. Immunol.* 5, 170.

Keşmir, C., and De Boer, R.J. (2003). A spatial model of germinal center reactions: cellular adhesion based sorting of B cells results in efficient affinity maturation. *J. Theor. Biol.* 222, 9–22.

Klein, F., Diskin, R., Scheid, J.F., Gaebler, C., Mouquet, H., Georgiev, I.S., Pancera, M., Zhou, T., Incesu, R.-B., Fu, B.Z., et al. (2013). Somatic mutations of the immunoglobulin framework are generally required for broad and potent HIV-1 neutralization. *Cell* 153, 126–138.

Kocks, C., and Rajewsky, K. (1988). Stepwise intraclonal maturation of antibody affinity through somatic hypermutation. *Proc. Natl. Acad. Sci. USA* 85, 8206–8210.

Kroese, F.G.M., Wubbena, A.S., Seijen, H.G., and Nieuwenhuis, P. (1987). Germinal centers develop oligoclonally. *Eur. J. Immunol.* 17, 1069–1072.

Kwong, P.D., Doyle, M.L., Casper, D.J., Cicala, C., Leavitt, S.A., Majeed, S., Steenbeke, T.D., Venturi, M., Chaiken, I., Fung, M., et al. (2002). HIV-1 evades antibody-mediated neutralization through conformational masking of receptor-binding sites. *Nature* 420, 678–682.

Kwong, P.D., Mascola, J.R., and Nabel, G.J. (2013). Broadly neutralizing antibodies and the search for an HIV-1 vaccine: the end of the beginning. *Nat. Rev. Immunol.* 13, 693–701.

Liao, H.-X., Lynch, R., Zhou, T., Gao, F., Alam, S.M., Boyd, S.D., Fire, A.Z., Roskin, K.M., Schramm, C.A., Zhang, Z., et al.; NISC Comparative Sequencing Program (2013). Co-evolution of a broadly neutralizing HIV-1 antibody and founder virus. *Nature* 496, 469–476.

Lyumkis, D., Julien, J.-P., de Val, N., Cupo, A., Potter, C.S., Klasse, P.J., Burton, D.R., Sanders, R.W., Moore, J.P., Carragher, B., et al. (2013). Cryo-EM structure of a fully glycosylated soluble cleaved HIV-1 envelope trimer. *Science* 342, 1484–1490.

Malherbe, D.C., Doria-Rose, N.A., Misher, L., Beckett, T., Puryear, W.B., Schuman, J.T., Kraft, Z., O'Malley, J., Mori, M., Srivastava, I., et al. (2011). Sequential immunization with a subtype B HIV-1 envelope quasispecies partially mimics the *in vivo* development of neutralizing antibodies. *J. Virol.* 85, 5262–5274.

Mascola, J.R., and Haynes, B.F. (2013). HIV-1 neutralizing antibodies: understanding nature's pathways. *Immunol. Rev.* 254, 225–244.

Meyer-Hermann, M. (2002). A mathematical model for the germinal center morphology and affinity maturation. *J. Theor. Biol.* 216, 273–300.

Meyer-Hermann, M.E., Maini, P.K., and Iber, D. (2006). An analysis of B cell selection mechanisms in germinal centers. *Math. Med. Biol.* 23, 255–277.

Nieuwenhuis, P., and Opstelten, D. (1984). Functional anatomy of germinal centers. *Am. J. Anat.* 170, 421–435.

Oprea, M., and Perelson, A.S. (1997). Somatic mutation leads to efficient affinity maturation when centrocytes recycle back to centroblasts. *J. Immunol.* 158, 5155–5162.

Pancera, M., Zhou, T., Druz, A., Georgiev, I.S., Soto, C., Gorman, J., Huang, J., Acharya, P., Chuang, G.Y., Ofek, G., et al. (2014). Structure and immune recognition of trimeric pre-fusion HIV-1 Env. *Nature* 514, 455–461.

Pantophlet, R., Wilson, I.A., and Burton, D.R. (2003). Hyperglycosylated mutants of human immunodeficiency virus (HIV) type 1 monomeric gp120 as novel antigens for HIV vaccine design. *J. Virol.* 77, 5889–5901.

Pissani, F., Malherbe, D.C., Robins, H., DeFilippis, V.R., Park, B., Sellhorn, G., Stamatatos, L., Overbaugh, J., and Haigwood, N.L. (2012). Motif-optimized subtype A HIV envelope-based DNA vaccines rapidly elicit neutralizing antibodies when delivered sequentially. *Vaccine* 30, 5519–5526.

Seaman, M.S., Janes, H., Hawkins, N., Grandpre, L.E., Devoy, C., Giri, A., Coffey, R.T., Harris, L., Wood, B., Daniels, M.G., et al. (2010). Tiered categorization of a diverse panel of HIV-1 Env pseudoviruses for assessment of neutralizing antibodies. *J. Virol.* 84, 1439–1452.

Selvarajah, S., Puffer, B., Pantophlet, R., Law, M., Doms, R.W., and Burton, D.R. (2005). Comparing antigenicity and immunogenicity of engineered gp120. *J. Virol.* 79, 12148–12163.

Shlomchik, M.J., and Weisel, F. (2012). Germinal centers. *Immunol. Rev.* 247, 5–10.

Shlomchik, M.J., Watts, P., Weigert, M.G., and Litwin, S. (1998). Clone: a Monte-Carlo computer simulation of B cell clonal expansion, somatic mutation, and antigen-driven selection. *Curr. Top. Microbiol. Immunol.* 229, 173–197.

Shulman, Z., Gitlin, A.D., Targ, S., Jankovic, M., Pasqual, G., Nussenzweig, M.C., and Victora, G.D. (2013). T follicular helper cell dynamics in germinal centers. *Science* 341, 673–677.

Swerdlin, N., Cohen, I.R., and Harel, D. (2008). The lymph node B cell immune response: dynamic analysis *in-silico*. *Proc. IEEE* 96, 1421–1443.

Thali, M., Furman, C., Ho, D.D., Robinson, J., Tilley, S., Pinter, A., and Sodroski, J. (1992). Discontinuous, conserved neutralization epitopes overlapping the CD4-binding region of human immunodeficiency virus type 1 gp120 envelope glycoprotein. *J. Virol.* 66, 5635–5641.

Victora, G.D., and Nussenzweig, M.C. (2012). Germinal centers. *Annu. Rev. Immunol.* 30, 429–457.

Victora, G.D., Schwickert, T.A., Fooksman, D.R., Kamphorst, A.O., Meyer-Hermann, M., Dustin, M.L., and Nussenzweig, M.C. (2010). Germinal center dynamics revealed by multiphoton microscopy with a photoactivatable fluorescent reporter. *Cell* 143, 592–605.

Wagner, S.D., Milstein, C., and Neuberger, M.S. (1995). Codon bias targets mutation. *Nature* 376, 732.

Walker, L.M., Huber, M., Doores, K.J., Falkowska, E., Pejchal, R., Julien, J.P., Wang, S.K., Ramos, A., Chan-Hui, P.Y., Moyle, M., et al.; Protocol G Principal Investigators (2011). Broad neutralization coverage of HIV by multiple highly potent antibodies. *Nature* 477, 466–470.

Wedemayer, G.J., Patten, P.A., Wang, L.H., Schultz, P.G., and Stevens, R.C. (1997). Structural insights into the evolution of an antibody combining site. *Science* 276, 1665–1669.

Wei, X., Decker, J.M., Wang, S., Hui, H., Kappes, J.C., Wu, X., Salazar-Gonzalez, J.F., Salazar, M.G., Kilby, J.M., Saag, M.S., et al. (2003). Antibody neutralization and escape by HIV-1. *Nature* 422, 307–312.

West, A.P., Jr., Scharf, L., Scheid, J.F., Klein, F., Bjorkman, P.J., and Nussenzweig, M.C. (2014). Structural insights on the role of antibodies in HIV-1 vaccine and therapy. *Cell* 156, 633–648.

Wyatt, R., Kwong, P.D., Desjardins, E., Sweet, R.W., Robinson, J., Hendrickson, W.A., and Sodroski, J.G. (1998). The antigenic structure of the HIV gp120 envelope glycoprotein. *Nature* 393, 705–711.

Ye, L., Zeng, R., Bai, Y., Roopenian, D.C., and Zhu, X. (2011). Efficient mucosal vaccination mediated by the neonatal Fc receptor. *Nat. Biotechnol.* 29, 158–163.

Zhang, J., and Shakhnovich, E.I. (2010). Optimality of mutation and selection in germinal centers. *PLoS Comput. Biol.* 6, e1000800.

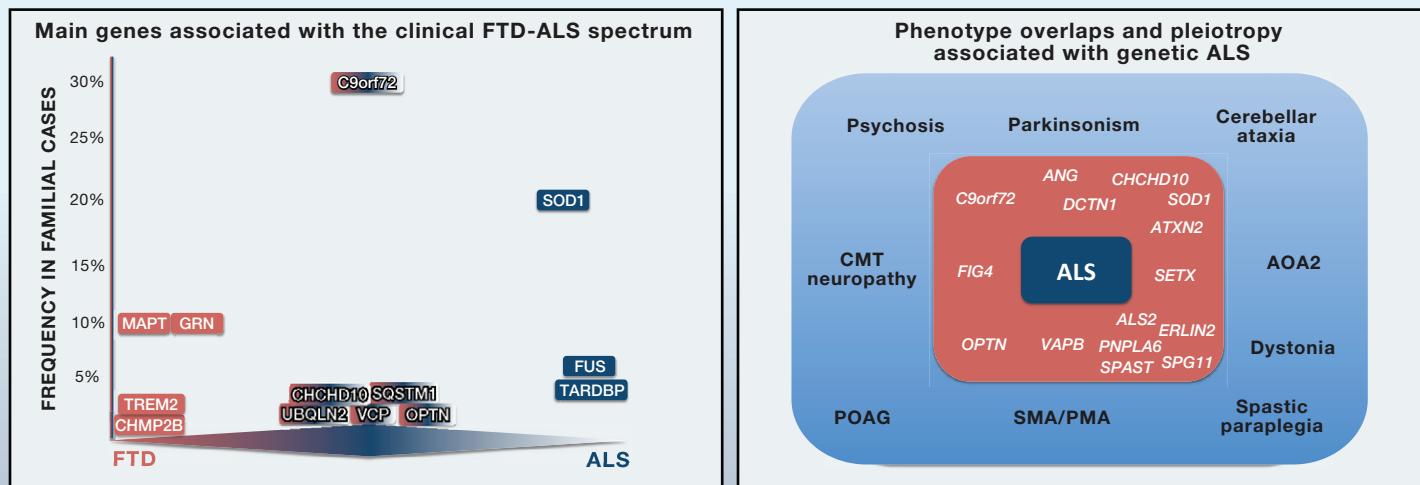
Zhou, T., Georgiev, I., Wu, X., Yang, Z.Y., Dai, K., Finzi, A., Kwon, Y.D., Scheid, J.F., Shi, W., Xu, L., et al. (2010). Structural basis for broad and potent neutralization of HIV-1 by antibody VRC01. *Science* 329, 811–817.

# SnapShot: Genetics of ALS and FTD

Cell

Rita Guerreiro, José Brás, and John Hardy

Department of Molecular Neuroscience, Institute of Neurology, University College London,  
Queen Square, London WC1N 1PJ, UK



Gene official symbol	Gene name	Location	Clinical presentation	Possible pathways / pathological biological processes
<b>MENTELIAN GENES</b>				
MAPT	Microtubule-associated protein tau	17q21.1	FTD	Toxic aggregation (defect in neuronal cytoskeleton)
GRN	Granulin	17q21.32	FTD*	Autophagy; lysosomal pathway; inflammation
TREM2	Triggering receptor expressed on myeloid cells 2	6p21.1	FTD	Inflammation
CHMP2B	Charged multivesicular body protein 2B	3p11.2	FTD	Autophagy; lysosomal pathway
C9orf72	Chromosome 9 open reading frame 72	9p21.2	FTD, ALS	Toxic RNA or repeat dipeptides aggregation
SQSTM1	Sequestosome 1	5q35	FTD, ALS*	Autophagy
UBQLN2	Ubiquilin 2	Xp11.21	FTD, ALS*	Autophagy
VCP	Valosin-containing protein	9p13.3	FTD, ALS*	Autophagy
OPTN	Optineurin	10p13	FTD/ALS*	Autophagy
SOD1	Superoxide dismutase 1, soluble	21q22.11	ALS	Toxic aggregation; free radical scavenger enzyme
FUS	FUS RNA binding protein	16p11.2	ALS	DNA/RNA metabolism
TARDBP	TAR DNA binding protein	1p36.22	ALS	DNA/RNA metabolism
CHCHD10	Coiled-coil-helix-coiled-coil-helix domain containing 10	22q11.23	FTD, ALS	Mitochondrial function
ALS2	Amyotrophic lateral sclerosis 2 (juvenile)	2q33.1	ALS*	Modulator for endosomal dynamics
SPG11	Spastic paraplegia 11 (autosomal recessive)	15q14	ALS*	DNA damage repair
SETX	Senataxin	9q34.13	ALS*	DNA/RNA processing
MATR3	Matrin 3	5q31.2	ALS*	DNA/RNA metabolism
ANG	Angiogenin, ribonuclease, RNase A family, 5	14q11.1-q11.2	ALS	Blood vessel formation
VAPB	VAMP (vesicle-associated membrane protein)-associated protein B and C	20q13.33	ALS*	Vesicle trafficking
PFN1	Profilin 1	17p13.3	ALS	Actin dynamics
TAF15	TAF15 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 68kDa	17q11.1-q11.2	ALS	RNA metabolism
HNRNPA1	Heterogeneous nuclear ribonucleoprotein A1	12q13.1	ALS, FTD*	RNA metabolism; direct interaction with TDP-43
HNRNPA2B1	Heterogeneous nuclear ribonucleoprotein A2/B1	7p15	ALS, FTD*	RNA metabolism; direct interaction with TDP-43
ERBB4	Erb-B2 receptor tyrosine kinase 4	2q33.3-q34	ALS	Dysregulation of the neuregulin-ErbB4 pathway
ARHGEF28	Rho guanine nucleotide exchange factor (GEF) 28	5q13.2	ALS	Interaction with low-molecular-weight neurofilament mRNA
DAO	D-amino-acid oxidase	12q24	ALS	
GLE1	GLE1 RNA export mediator	9q34.11	ALS	RNA metabolism
SIGMAR1	Sigma non-opioid intracellular receptor 1	9p13.3	ALS, FTD	Endoplasmic reticulum lipid rafts
ERL1N2	ER lipid raft associated 2	8p11.2	ALS*	Endoplasmic reticulum lipid rafts
PNPLA6	Patatin-like phospholipase domain containing 6	19p13.2	ALS*	Neurite outgrowth and process elongation
PRKAR1B	Protein kinase, cAMP-dependent, regulatory, type I, beta	7p22	FTD*	Regulation of metabolism, ion transport, and gene transcription
DCTN1	Dynactin 1	2p13	ALS*	Vesicle trafficking
<b>RISK OR MODIFIER GENES</b>				
TMEM106B	Transmembrane protein	7p21.3	FTD	Regulation of lysosomal function and progranulin pathways
ATXN2	Ataxin 2 (susceptibility)	12q24.1	FTD/ALS	Modifier of tdp-43 toxicity
UNC13A/ KCNN1	Unc-13 homolog A ( <i>C. elegans</i> )/potassium channel, calcium-activated intermediate/small conductance subfamily N alpha, member 1	19p13.11 /19p13.1	ALS	Neurotransmitter release/neuronal excitability
ZNF512B	Zinc finger protein 512B	20q13.33	ALS	

# SnapShot: Genetics of ALS and FTD

Rita Guerreiro, José Brás, and John Hardy

Department of Molecular Neuroscience, Institute of Neurology, University College London,  
Queen Square, London WC1N 1PJ, UK

Cell

Frontotemporal dementia (FTD) and amyotrophic lateral sclerosis (ALS) are considered to be part of a spectrum. Clinically, FTD patients present with dementia frequently characterized by behavioral and speech problems. ALS patients exhibit alterations of voluntary movements caused by degeneration of motor neurons. Both syndromes can be present within the same family or even in the same person. The genetic findings for both diseases also support the existence of a continuum, with mutations in the same genes being found in patients with FTD, ALS, or FTD/ALS.

The figure on the left represents the distribution of genes according to their associated phenotype and mutation frequency in familial cases (this can vary between cohorts or populations).

Some cases of mutations in genes represented in the extremes of the graph (associated with either ALS or FTD) have also been described as associated with the other clinical phenotypes. For example, mutations in *TARDBP* that are the cause of ALS have, in rare occasions, been described in cases presenting with FTD. The genetic evidence for these associations is weak (mainly lacking segregation of the variants with the phenotypes or strong evidence of pathogenicity), and for these reasons, such genes are represented as the extremes of the FTD-ALS continuum.

Genetic ALS is usually not only associated with motor systems degeneration but also, very frequently, occurs in combination with degenerative processes like parkinsonism or ataxia. In fact, very few ALS genes have been associated with a strict motor-only phenotype, and genes commonly associated with other diseases have sometimes been found to be the cause of ALS. This was the case for *FIG4* and *SPAST* mutations that have been originally associated with Charcot-Marie-Tooth neuropathy type 4J (CMT4J) and autosomal-dominant spastic paraparesis-4 (SPG4), respectively, and have been subsequently found to be mutated in ALS cases.

These overlaps between different phenotypes and genes found to be mutated either in ALS with additional features or in other diseases and ALS are represented in the right panel.

One FTD gene (*PRKAR1B*) and several ALS genes (*ARHGEF28*, for example) have been, so far, only reported in a single family (with different levels of segregation) or have not yet been independently replicated. *PRKAR1B* was identified in an FTD-like family with a syndrome considered to be novel. The family presented dementia and/or parkinsonism in 12 affected individuals and a unique neuropathology displaying abundant neuronal inclusions by hematoxylin and eosin staining throughout the brain with immunoreactivity for intermediate filaments. Some ALS genes have conflicting results with variants described in controls and uncertain pathogenicity attributed to different mutations (*ANG*, for example).

Several genes implicated in FTD and ALS have possible functions within the same biological pathways such as autophagy and the lysosomal pathway, maintenance of neuronal cytoskeleton, and DNA/RNA metabolism. In some cases, genes also share similar domains like the RNA binding genes with prion-like domains *HNRNPA1*, *HNRNPA2B1*, *TARDBP*, *FUS*, *TAF15*, and *EWSR1*. In fact, *TAF15* and *EWSR1* were screened for the identification of variants because of their commonalities with *FUS* and *TARDBP*. In both genes, variants were found in patients with sporadic ALS, and weak familial segregation has only been demonstrated for *TAF15*.

The use of next-generation sequencing technologies has recently allowed the identification of rare variants in *TUBA4A* to be associated with ALS, and the analyses of de novo mutations in sporadic ALS trios have identified mutations in *SS18L1*. Again, the approaches used did not allow for segregation analyses and, for this reason, even though the variants were associated with familial ALS, the genes are not included as a Mendelian gene for ALS in the table.

Other genes have been reported as susceptibility or modifiers of FTD and/or ALS with different levels of evidence. These include, in addition to the ones represented in the table, *BTNL2/HLA-DRA/HLA-DRB5*, *RAB38/CTSC*, *CHRNA3/4*, *CHRN4*, *PRPH*, *CHGB*, *NEFH*, *VEGF*, *HFE*, *PON1/2/3*, *DPP6*, *ELP3*, *ITPR2*, *FGGY*, *CHGB*, *DPYSL3*, *KIFAP3*, *EPHA3/4*, *PPARGC1A*, *APOE*, *MAOB*, *CX3CR1*, and *SMN1*.

The loci identified by genome-wide association studies are named according to the closest gene to the most significant hit at each locus. It is possible that the gene involved in FTD and/or ALS is not the closest but is in linkage disequilibrium with the true hit. This has been recently shown to be the case for *UNC13A*, where *KCNN1* (a nearby gene) was implicated by eQTL analysis.

An asterisk (\*) indicates genes for which mutations may present additional phenotypes or be the cause of different clinical syndromes. For example, loss-of-function heterozygous mutations in *GRN* cause FTD, whereas homozygous mutations in the same gene have recently been reported to cause neuronal ceroid lipofuscinosis-11 (CLN11); *VCP* mutations are known to cause ALS with or without FTD and inclusion body myopathy with Paget disease of bone and frontotemporal dementia (IBMPFD).

## ABBREVIATIONS

CMT neuropathy, Charcot-Marie-Tooth neuropathy; POAG, primary open angle glaucoma; SMA/PMA, spinal muscular atrophy/progressive muscular atrophy; AOA2, Ataxia with oculomotor apraxia type 2.

## ACKNOWLEDGMENTS

Research studies in the authors' lab are mainly supported by the Alzheimer's Research UK (ARUK), including a Fellowship to RG; by a Fellowship from Alzheimer's Society to J.B.; by the Wellcome Trust/MRC Joint Call in Neurodegeneration award (WT089698) to the UK Parkinson's Disease Consortium, whose members are from the UCL/Institute of Neurology, the University of Sheffield, and the MRC Protein Phosphorylation Unit at the University of Dundee; and by an anonymous donor.

## REFERENCES

Abel, O., Powell, J.F., Andersen, P.M., and Al-Chalabi, A. (2012). *Hum. Mutat.* 33, 1345–1351.

Andersen, P.M., and Al-Chalabi, A. (2011). *Nat. Rev. Neurol.* 7, 603–615.

Hardy, J., and Rogeava, E. (2014). *Exp. Neurol.* 262 Pt B, 75–83.

Marangi, G., and Traynor, B.J. (2014). *Brain Res.* Published online October 12, 2014. <http://dx.doi.org/10.1016/j.brainres.2014.10.009>.

Ramasamy, A., Trabzuni, D., Guelfi, S., Varghese, V., Smith, C., Walker, R., De, T., Coin, L., de Silva, R., Cookson, M.R., et al.; UK Brain Expression Consortium; North American Brain Expression Consortium (2014). *Nat. Neurosci.* 17, 1418–1428.

Smith, B.N., Ticozzi, N., Fallini, C., Gkazi, A.S., Topp, S., Kenna, K.P., Scotter, E.L., Kost, J., Keagle, P., Miller, J.W., et al.; SLAGEN Consortium (2014). *Neuron* 84, 324–331.

Wong, T.H., Chiu, W.Z., Breedveld, G.J., Li, K.W., Verkerk, A.J., Hondius, D., Hukema, R.K., Seelaar, H., Frick, P., Severijnen, L.A., et al. (2014). *Brain* 137, 1361–1373.